Harvey DI, Leyboume SJ, Sollis R, Taylor AMR.

Tests for explosive financial bubbles in the presence of non-stationary volatility.

Journal of Empirical Finance 2016, 38(Part B), 548-574.

DOI link to article:

https://doi.org/10.1016/j.jempfin.2015.09.002

Date deposited:

22/02/2018

Embargo release date:

25 March 2017

Tests for Explosive Financial Bubbles in the Presence of Non-stationary Volatility*

David I. Harvey[a], Stephen J. Leybourne[a], Robert Sollis[b] and A.M. Robert Taylor[c]

[a]School of Economics, University of Nottingham

[b]Newcastle University Business School

[c]Essex Business School, University of Essex

May 15, 2015

Abstract

This paper studies the impact of permanent volatility shifts in the innovation process on the performance of the test for explosive financial bubbles based on recursive right-tailed Dickey-Fuller-type unit root tests proposed by Phillips, Wu and Yu (2011). We show that, in this situation, their supremum-based test has a non-pivotal limit distribution under the unit root null, and can be quite severely over-sized, thereby giving rise to spurious indications of explosive behaviour. We investigate the performance of a wild bootstrap implementation of their test procedure for this problem, and show it is effective in controlling size, both asymptotically and in finite samples, yet does not sacrifice power relative to an (infeasible) size-adjusted version of their test, even when the shocks are homoskedastic. We also discuss an empirical application involving commodity price time series and find considerably less emphatic evidence for the presence of speculative bubbles in these data when using our proposed wild bootstrap implementation of the Phillips, Wu and Yu (2011) test.

Keywords: Rational bubble; Explosive autoregression; Non-stationary volatility; Right-tailed unit root testing.

JEL Classification: C22; C12; G14.

1 Introduction

In seminal research on the presence of explosive rational asset price bubbles in stock prices, Diba and Grossman (1988) highlight the usefulness of unit root tests for detecting such bubbles. They note that if the bubble component of the stock price evolves as an explosive autoregressive

*We are grateful to two anonymous referees and participants at the conference 'Recent Developments in Financial Econometrics and Empirical Finance', held at the University of Essex on June 12th and 13th 2014, for helpful comments. Correspondence to: Robert Taylor, Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. Email: `rtaylor@essex.ac.uk`

process then, since an explosive autoregressive process cannot be differenced to stationarity, a finding of non-stationarity for the price and dividend series when the series are in levels but stationarity when the series are in first differences is indicative that an explosive rational bubble does not exist. Consequently, Diba and Grossman (1988) proposed testing the no bubble hypothesis by applying orthodox left-tailed unit root tests to the price and dividend series in levels and first-differenced forms. More recently, researchers in this area have focused on testing for explosive autoregressive behaviour directly via the application of right-tailed Dickey-Fuller [DF] tests. Phillips *et al.* (2011) [PWY] were the first to employ this approach. They suggest a test procedure for detecting explosive rational bubbles in stock prices based on the supremum of a set of forward recursive right-tailed DF test statistics applied to the price and dividend series in levels only. If the test finds explosive autoregressive behaviour for the prices (but not for the dividends), this indicates the presence of an explosive rational bubble.

More generally, it should also be recognised that whilst finding evidence of explosive autoregressive behaviour in prices that is not driven by fundamentals suggests that explosive rational price bubbles exist, there are other intuitively sensible non-bubble explanations for this pattern of results. For example, an alternative perspective is that non-explosive changes in fundamentals could quite feasibly lead to structural changes in asset prices to new equilibrium levels, and the transition to new equilibria might appear to be explosive. This argument is consistent with earlier research on the 1990s stock market boom which argues that fundamentals-based explanations such as lower expected future real discount rates and higher expected real dividend (earnings) growth cannot be ruled out as an explanation for the apparently explosive prices (see, for example, Balke and Wohar, 2001). Notwithstanding such considerations, in this paper we use the terminology "bubble" interchangeably with evidence for explosive autoregressive characteristics in prices, implicitly assuming non-explosive fundamentals.

The PWY test is simple to apply and Monte Carlo simulations reported in PWY show that it has very good finite sample power to detect an explosive asset price bubble. PWY apply their proposed testing procedure to the Nasdaq Composite stock price index and dividend index between February 1973 and June 2005. Their results indicate the presence of an explosive rational bubble, and subsequent application of their dating methodology suggests that the bubble began in mid-1995. The PWY test and related tests have become popular with applied researchers investigating the presence or otherwise of speculative bubbles in various different financial price series data. Gilbert (2010) employs the PWY test to investigate for the presence of speculative bubbles in commodities futures prices over 2000-2009, finding evidence of bubbles in the copper, nickel and crude oil markets. Homm and Breitung (2012) apply the PWY test and a related Chow-type test to stock price, commodity price and house price data, finding evidence of bubbles in many of the series examined. Bettendorf and Chen (2013) use the PWY test to look for explosive bubbles in the sterling-US dollar nominal exchange rate. They find statistically significant evidence of explosive behaviour in the nominal exchange rate and this appears to be driven by explosive behaviour in the relevant price index ratio for traded goods.

Our focus in this paper is on the performance of the PWY test to detect explosive autoregressive behaviour in cases where the volatility of the innovation process is subject to non-stationarity, a leading example of which is where structural breaks occur in the unconditional variance of the innovation process. A growing number of applied studies have found strong evidence of structural breaks in the unconditional variance of asset returns, often with the breaks linked to major financial and macroeconomic crises such as the 1970s oil price shocks, the East Asian currency crisis in the late-1990s, the dot-com crash in 2001 and the recent global financial crisis in 2007-2009. Indeed in a number of these studies very large structural breaks have been detected; for example, Rapach *et al.* (2008) and McMillan and Wohar (2011) detect breaks in the unconditional variance of the returns of some major stock market indices and sectoral stock price indices, finding that the unconditional variance in some sub-samples can be larger than that in other sub-samples by a factor of about ten. For commodity returns, both Calvo-Gonzalez *et al.* (2010) and Vivian and Wohar (2012) find statistically significant evidence of structural breaks in unconditional volatility. Needless to say, volatility changes in innovations to price series processes could be induced by the presence of a speculative bubble, but equally it could be the case that changes in volatility occur without an explosive bubble period occurring. It is therefore critically important to have available a reliable method for detecting an explosive period in a series that is robust to the potential presence of non-stationary volatility. This is particularly important if the evidence is intended to inform future monetary policy.

A key feature of the PWY test is that, as with the orthodox DF test, it assumes that the unconditional variance of the innovation process is stationary under both the unit root null hypothesis and the explosive alternative hypothesis. If the PWY test is applied to prices, this assumption implies that when a bubble does not exist, the unconditional variance of the innovations does not undergo permanent shifts of any form. Thus, if there was, say, a major financial/macroeconomic crisis that increased unconditional volatility, then so the PWY test applied to the price series would be inherently misspecified. If such a crisis was not preceded by an asset price bubble, then market efficiency arguments would suggest that the price series will follow a unit root process; that is, the null hypothesis associated with the PWY test is true, but the volatility break could have an impact on the size properties of the PWY test. The most serious consequence of this would be spurious rejections of the no bubble hypothesis, indicating the presence of a bubble when one does not actually exist.

We first analyse the asymptotic properties of the PWY test statistic when non-stationary volatility is allowed for in the innovations. We show that the limiting distribution of the PWY statistic depends on nuisance parameters derived from the pattern of heteroskedasticity present in the innovations, and that this holds under both the null and local alternatives. We quantify these effects for a variety of non-stationary volatility processes including single and double breaks in volatility and trending volatility. These results show that the PWY test can be badly over-sized for plausible models of non-stationary volatility and, as a result, spuriously reject the unit root null hypothesis in favour of explosive behaviour.

In response to the inference problem we identify with the standard PWY test, we consider a simple solution that restores correct asymptotic size. In particular, we propose use of a wild bootstrap scheme, applied to the first differences of the data, in order to replicate in the bootstrap data the pattern of non-stationary volatility present in the original innovations.[1] Etienne *et al.* (2014, 2015) apply a similar wild bootstrap approach to the Phillips *et al.* (2014) extension of the PWY procedure in their analysis of food commodity markets, but do not develop the theoretical properties of their procedure.[2] We show that the wild bootstrap analogues of the PWY statistic share the same (first order) limiting null distribution as the original statistics within a broad class of non-stationary volatility processes. Hence, asymptotic inference is rendered robust to the potential presence of non-stationary volatility in the innovations without requiring the practitioner to specify a parametric model for the volatility process. Importantly, we also demonstrate that our proposed bootstrap PWY test achieves the asymptotic local power function of an (infeasibly) size-corrected implementation of the original PWY statistic, under locally explosive alternatives. Under fixed magnitude explosive alternatives our bootstrap PWY test is shown to be consistent, although its finite sample power may no longer match that of a size-corrected original PWY statistic. To this end, we also consider a second bootstrap PWY test procedure that achieves the power of the size-corrected original PWY statistic under both locally and fixed explosive alternatives. This is based on fitting a model of the explosive regime to the data using the Bayesian information criterion [BIC] based model selection procedure of Harvey *et al.* (2015) [HLS]. We find there is rather little to choose between the powers (and sizes) of the simple first differences and model-based bootstrap procedures, which we take as supporting evidence that the simpler procedure is more than adequate for practical implementation.

The paper is organised along the following lines. In section 2 we introduce our reference data generation process (DGP), which categorises the type of explosive behaviour we consider, and details the class of non-stationary volatility within which we work. Section 3 outlines the standard PWY test for detecting explosivity. In section 4 we establish the large sample behaviour of the PWY statistic under both the unit root null and locally explosive alternatives when non-stationary volatility is present. Here we quantify the impact on the asymptotic size of the PWY test of a number of empirically relevant models of non-stationary volatility. Our simple wild bootstrap procedure is proposed in section 5 and its asymptotic properties are established under both the null and local alternatives. The behaviour of our proposed wild bootstrap procedure under a fixed magnitude explosive alternative is also examined, and the model-based alternate procedure is introduced. In section 6 we examine the finite sample

[1] Application of the wild bootstrap has been successfully used to robustify left-tailed unit root tests (that reject in favour of a stationary alternative) to non-stationary volatility; see, for exanple, Cavaliere and Taylor (2008a, 2009a).

[2] Gutierrez (2013) also uses a bootstrap version of PWY, but adopts a residual-based re-sampling scheme, which would not deliver robustness to non-stationary volatility.

size and power of the standard PWY test and both bootstrap variants under non-stationary volatility. Section 7 discusses an empirical application involving several commodity price time series. Section 8 concludes.

In the following: $1(.)$ denotes the indicator function; $\lfloor\cdot\rfloor$ denotes the integer part; $x := y$ ($x =: y$) indicates that x is defined by y (y is defined by x); $\xrightarrow{w}$ denotes weak convergence, $\xrightarrow{p}$ convergence in probability, and $\xrightarrow{w}_p$ weak convergence in probability, in each case as the sample size diverges; finally, $\mathcal{C} = C[0,1]$ denotes the space of continuous processes on $[0,1]$, and $\mathcal{D} = D[0,1]$ the space of right continuous with left limit (càdlàg) processes on $[0,1]$.

2 The Heteroskedastic Bubble Model

We will consider the time series process $\{y_t\}$ generated according to the following DGP,

$$y_t = \mu + u_t \tag{1}$$

$$u_t = \begin{cases} u_{t-1} + \varepsilon_t, & t = 2, ..., \lfloor \tau_{1,0} T \rfloor, \\ (1+\delta_{1,T})u_{t-1} + \varepsilon_t, & t = \lfloor \tau_{1,0} T \rfloor + 1, ..., \lfloor \tau_{2,0} T \rfloor, \\ (1-\delta_{2,T})u_{t-1} + \varepsilon_t, & t = \lfloor \tau_{2,0} T \rfloor + 1, ..., \lfloor \tau_{3,0} T \rfloor, \\ u_{t-1} + \varepsilon_t, & t = \lfloor \tau_{3,0} T \rfloor + 1, ..., T \end{cases} \tag{2}$$

where $\delta_{1,T} \geq 0$ and $\delta_{2,T} \geq 0$. We assume that the initial condition u_1 is such that $u_1 = o_p(T^{1/2})$, while the innovation process $\{\varepsilon_t\}$ satisfies the following assumption:

Assumption 1 *Let $\varepsilon_t = \sigma_t z_t$ where $z_t \sim IID(0,1)$ with $E|z_t|^r < K < \infty$ for some $r \geq 4$. The volatility term σ_t satisfies $\sigma_t = \omega(t/T)$, where $\omega(\cdot) \in \mathcal{D}$ is non-stochastic and strictly positive. For $t \leq 0$, $\sigma_t \leq \breve{\sigma} < \infty$.*

When $\delta_{1,T} > 0$, y_t follows a unit root process up to time $\lfloor \tau_{1,0} T \rfloor$, after which point it displays explosive autoregressive behaviour over the period $t = \lfloor \tau_{1,0} T \rfloor + 1, ..., \lfloor \tau_{2,0} T \rfloor$. At the termination of the bubble period, the DGP in (1)-(2) admits two possibilities: if $\delta_{2,T} = 0$, y_t reverts to unit root dynamics directly, while if $\delta_{2,T} > 0$, the unit root dynamics resume after an interim stationary regime over the time period $t = \lfloor \tau_{2,0} T \rfloor + 1, ..., \lfloor \tau_{3,0} T \rfloor$. This specification follows HLS and provides a model of a crash regime, where the mean-reverting stationary behaviour in this regime acts to "offset" the explosive period to some extent.[3] The magnitude of $\delta_{2,T}$ and the duration of the collapse regime ($\lfloor \tau_{3,0} T \rfloor - \lfloor \tau_{2,0} T \rfloor$) control the rapidity and extent to which a collapse occurs. This approach offers a flexible way of modelling a range of potential price corrections that might be expected when a bubble terminates, from relatively slow gradual adjustments in the price level to more rapid crashes; at the extreme, if a collapse to a lower level occurs instantaneously, the stationary regime acts as an approximation, although typically more gradual collapses are observed in practice as agents adjust their behaviour over

[3]Subsequent to the first version of HLS, Phillips and Shi (2014) adopt a similar model for a gradually collapsing regime.

a number of time periods. The DGP in (1)-(2) also admits a bubble (or collapse) regime continuing to the end of the sample period, on letting $\tau_{2,0} = 1$ (or $\tau_{3,0} = 1$). When $\delta_{1,T} = 0$, no explosive regime is present in the data, and we assume $\delta_{2,T} = 0$ also in this case, so that collapse regimes do not occur without a prior bubble.

The null hypothesis, H_0, is that no bubble is present in the series and y_t follows a unit root process throughout the sample period, i.e. $H_0 : \delta_{1,T} = 0$ (and hence $\delta_{2,T} = 0$). The alternative hypothesis is given by $H_1 : \delta_{1,T} > 0$, and corresponds to the case where a bubble is present in the series, which either runs to the end of the sample (if $\tau_{2,0} = 1$), or terminates in-sample, either with or without a subsequent collapse regime depending on whether $\delta_{2,T} = 0$ or $\delta_{2,T} > 0$.

Assumption 1 coincides with the set of conditions adopted in Cavaliere and Taylor (2007, 2008a) for the case where ε_t is serially uncorrelated.[4] The key assumption for the purposes of this paper is that the innovation variance is non-stochastic, bounded and displays a countable number of jumps. A detailed discussion of the class of variance processes allowed is given in Cavaliere and Taylor (2007); this includes variance processes displaying (possibly) multiple one-time volatility shifts (which need not be located at the same point in the sample as the putative regimes associated with bubble behaviour), polynomially (possibly piecewise) trending volatility and smooth transition variance breaks, among others. The conventional homoskedasticity assumption, that $\sigma_t = \sigma$ for all t, is also permitted, since here $\omega(s) = \sigma$ for all s. Assumption 1 requires that the volatility process is non-stochastic and that z_t is an IID sequence. These restrictions are placed in order to simplify our analysis but can be weakened, without affecting the main results of the paper, to allow for cases where $\omega(\cdot)$ is stochastic and independent of z_t and where z_t is a martingale difference sequence satisfying certain moment conditions; see Cavaliere and Taylor (2009b) for further details.

A quantity which will play a key role in what follows is given by the following function in $\mathcal{C}$, known as the *variance profile* of the process:

$$\eta(s) := \left(\int_0^1 \omega(h)^2 \, dh \right)^{-1} \int_0^s \omega(h)^2 \, dh.$$

Observe that the variance profile satisfies $\eta(s) = s$ under homoskedasticity while it deviates from s in the presence of heteroskedasticity. Notice also that the quantity $\overline{\omega}^2 := \int_0^1 \omega(h)^2 \, dh$ is equal to the limit of $T^{-1} \sum_{t=1}^T \sigma_t^2$, and may therefore be interpreted as the (asymptotic) average innovation variance. We will also make use of the invariance principle from Theorem 1(i) of Cavaliere and Taylor (2007), which establishes that

$$T^{-1/2} \sum_{j=1}^{\lfloor \cdot T \rfloor} \varepsilon_j \xrightarrow{w} \overline{\omega} W^{\eta}(\cdot)$$

where the process $W^{\eta}(r) := \int_0^r dW(\eta(s))$, $W(r)$ denoting a standard Brownian motion on $[0, 1]$, is known as a *variance-transformed* Brownian motion, i.e. a Brownian motion under a modification of the time domain; see, for example, Davidson (1994).

[4] Generalisations to allow for serial correlation in ε_t will be discussed in section 5.4 below.

3 The PWY Test Procedure

In this section we briefly review the PWY procedure for detecting (and date stamping) explosive bubbles, together with an alternative date stamping procedure developed recently in HLS. All of the material reviewed in this section is based on the assumption that the innovation process, ε_t in DGP (1)-(2) is homoskedastic; that is, Assumption 1 with $\sigma_t = \sigma$ for all t.

The PWY statistic is used to test H_0 against H_1 in the context of (1)-(2), the alternative being that y_t behaves as an explosive AR(1) process for at least some sub-period of the sample. In this context, and in the absence of knowledge concerning the timing of any potential explosive behaviour, and the precise nature of any collapse behaviour, PWY propose a test based on the supremum of recursive right-tailed DF tests. Specifically, for non-serially correlated ε_t, the PWY statistic is given by

$$PWY := \sup_{\tau \in [\tau_0, 1]} DF_\tau$$

where DF_τ denotes a standard DF statistic, that is the t-ratio for $\hat{\phi}_\tau$ in the fitted ordinary least squares [OLS] regression

$$\Delta y_t = \hat{\alpha}_\tau + \hat{\phi}_\tau y_{t-1} + \hat{\varepsilon}_{t,\tau} \tag{3}$$

calculated over the sub-sample period $t = 1, ..., \lfloor \tau T \rfloor$; that is,

$$DF_\tau := \frac{\hat{\phi}_\tau}{\sqrt{\hat{\sigma}_\tau^2 / \sum_{t=2}^{\lfloor \tau T \rfloor} (y_{t-1} - \bar{y}_\tau)^2}}$$

where $\bar{y}_\tau = (\lfloor \tau T \rfloor - 1)^{-1} \sum_{t=2}^{\lfloor \tau T \rfloor} y_{t-1}$ and $\hat{\sigma}_\tau^2 = (\lfloor \tau T \rfloor - 3)^{-1} \sum_{t=2}^{\lfloor \tau T \rfloor} \hat{\varepsilon}_{t,\tau}^2$. The PWY statistic is therefore the supremum of a sequence of forward recursive DF statistics with minimum sample length $\lfloor \tau_0 T \rfloor$. The choice of τ_0 involves a trade-off between incorporating sufficient initial observations in the first recursive sample while ensuring that the possibility of a short early bubble is accommodated; in what follows we follow PWY and set $\tau_0 = 0.1$. The PWY test rejects for large values of the PWY statistic with selected critical values given in Table 1 of PWY. They derive the limiting null distribution of the PWY statistic and show that the associated test is consistent against H_1.

If the PWY test signals a rejection then one would naturally wish to date stamp the emergence and, where appropriate, collapse of the period of exuberance. In the context of a simplified version of DGP (1)-(2) where the collapse (should one occur) happens instantaneously, PWY propose a procedure to do this based on the sequence of forward recursive DF statistics used in calculating the PWY supremum statistic. Specifically, they suggest locating the origin and conclusion of the explosive regime, by matching the time series of the recursive test statistic DF_τ, with $\tau \in [\tau_0, 1]$, against right-tail critical values. That is, they propose the following estimates of $\tau_{1,0}$ and $\tau_{2,0}$: $\hat{\tau}_{1,0} := \inf_{\tau \geq \tau_0} \{\tau : DF_\tau > cv_T(\tau)\}$ and $\hat{\tau}_{2,0} := \inf_{\tau \geq \hat{\tau}_{1,0}} \{\tau : DF_\tau < cv_T(\tau)\}$ (possibly subject to some minimum bubble duration requirement), where $cv_T(\tau)$ is a critical value that needs to diverge to infinity as T diverges to ensure consistent estimates under H_1; appropriate settings to achieve this are discussed in detail in PWY.

In the context of the DGP in (1)-(2) where the possibility of a crash regime is retained, HLS suggest an alternative date stamping approach based on BIC model selection. Briefly, this procedure considers four possible DGPs arising from (1)-(2) under H_1. Namely,

DGP 1: $\delta_{1,T} > 0$, $0 < \tau_{1,0} < 1$, $\tau_{2,0} = 1$
(unit root, then bubble to sample end)

DGP 2: $\delta_{1,T} > 0$, $\delta_{2,T} = 0$, $0 < \tau_{1,0} < \tau_{2,0} < 1$
(unit root, then bubble, then unit root to sample end)

DGP 3: $\delta_{1,T} > 0$, $\delta_{2,T} > 0$, $0 < \tau_{1,0} < \tau_{2,0} < 1$, $\tau_{3,0} = 1$
(unit root, then bubble, then collapse to sample end)

DGP 4: $\delta_{1,T} > 0$, $\delta_{2,T} > 0$, $0 < \tau_{1,0} < \tau_{2,0} < \tau_{3,0} < 1$
(unit root, then bubble, then collapse, then unit root to sample end).

Four corresponding models are then fitted to capture each possible DGP:

Model 1: $\Delta y_t = \hat{\alpha}_1 D_t(\tau_1, 1) + \hat{\beta}_1 D_t(\tau_1, 1) y_{t-1} + \hat{\varepsilon}_{1t}$
Model 2: $\Delta y_t = \hat{\alpha}_1 D_t(\tau_1, \tau_2) + \hat{\beta}_1 D_t(\tau_1, \tau_2) y_{t-1} + \hat{\varepsilon}_{2t}$
Model 3: $\Delta y_t = \hat{\alpha}_1 D_t(\tau_1, \tau_2) + \hat{\alpha}_2 D_t(\tau_2, 1) + \hat{\beta}_1 D_t(\tau_1, \tau_2) y_{t-1} + \hat{\beta}_2 D_t(\tau_2, 1) y_{t-1} + \hat{\varepsilon}_{3t}$
Model 4: $\Delta y_t = \hat{\alpha}_1 D_t(\tau_1, \tau_2) + \hat{\alpha}_2 D_t(\tau_2, \tau_3) + \hat{\beta}_1 D_t(\tau_1, \tau_2) y_{t-1} + \hat{\beta}_2 D_t(\tau_2, \tau_3) y_{t-1} + \hat{\varepsilon}_{4t}$

where $D_t(a, b) = 1(\lfloor aT \rfloor < t \leq \lfloor bT \rfloor)$. In each case the change-point estimators are obtained by minimising the sum of squared residuals across all permitted possibilities (subject to $\tau_1 \geq s$, $\tau_2 - \tau_1 \geq s$, $\tau_3 - \tau_2 \geq s/2$; we use $s = 0.1$ throughout), and subject to the requirement that $y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_1 T \rfloor}$ and $y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_3 T \rfloor}$, ensuring that the period from τ_1 to τ_2 is associated with a (putative) upward explosive regime, and τ_2 to τ_3 associates with a downward stationary collapse regime. In each model the final regime is permitted to be of any length, providing a smooth segue from one model to another; see section 5 of HLS for full details. A choice between these alternative estimated models is then made on the basis of the usual BIC, penalising both the number of estimated dummy variable parameters and the number of estimated regime change dates. HLS show that in the limit the model corresponding to the true DGP is selected with probability one under H_1 when the bubble (and collapse) parameters are of fixed magnitudes. Moreover, the change-point estimators associated with the selected model are such that $\lfloor \hat{\tau}_i T \rfloor - \lfloor \tau_{i,0} T \rfloor \xrightarrow{p} 0$, that is, the actual *dates* of the start and end of the bubble (and collapse) periods are consistently estimated.

4 Asymptotic Behaviour of the PWY Test

In this section we analytically investigate the impact of non-stationary volatility of the form given in Assumption 1 on the large sample behaviour of the PWY statistic under both H_0 and H_1. Under H_1 we consider local-to-unit root settings for the explosive and stationary regime parameters, i.e. $\delta_{i,T} = c_i T^{-1}$, $i = 1, 2$, $c_1 > 0$, $c_2 \geq 0$, the scalings by T^{-1} providing the appropriate Pitman drifts for the DGP in (1)-(2).

In Theorem 1 we now provide the asymptotic distribution of the PWY statistic under H_1, the corresponding result under H_0 being obtained as a special case thereof.

Theorem 1. *Let* $\{y_t\}$ *be generated according to* (1)-(2) *under Assumption 1 and with* $\delta_{i,T} = c_i T^{-1}$, $c_i \geq 0$, $i = 1, 2$. *Then,*

$$PWY \xrightarrow{w} \overline{\omega} \sup_{\tau \in [\tau_0, 1]} L^{\eta}_{c_1,c_2}(\tau) =: S^{\eta}_{c_1,c_2} \tag{4}$$

where

$$L^{\eta}_{c_1,c_2}(\tau) := \frac{1}{\sqrt{\tau^{-1}\int_0^{\tau} \omega(h)^2 \, dh}} \frac{\int_0^{\tau} \tilde{K}^{\eta}_{c_1,c_2}(r) dK^{\eta}_{c_1,c_2}(r)}{\sqrt{\int_0^{\tau} \tilde{K}^{\eta}_{c_1,c_2}(r)^2 dr}}$$

and

$$\tilde{K}^{\eta}_{c_1,c_2}(r) := K^{\eta}_{c_1,c_2}(r) - \tfrac{1}{\tau}\int_0^{\tau} K^{\eta}_{c_1,c_2}(s) ds$$

with

$$K^{\eta}_{c_1,c_2}(r) := \begin{cases} W^{\eta}(r) & r \leq \tau_{1,0} \\ e^{(r-\tau_{1,0})c_1} W^{\eta}(\tau_{1,0}) + \int_{\tau_{1,0}}^{r} e^{(r-s)c_1} dW^{\eta}(s) & \tau_{1,0} < r \leq \tau_{2,0} \\ e^{-(r-\tau_{2,0})c_2}\left\{ e^{(\tau_{2,0}-\tau_{1,0})c_1} W^{\eta}(\tau_{1,0}) + \int_{\tau_{1,0}}^{\tau_{2,0}} e^{(\tau_{2,0}-s)c_1} dW^{\eta}(s)\right\} & \\ \quad + \int_{\tau_{2,0}}^{r} e^{-(r-s)c_2} dW^{\eta}(s) & \tau_{2,0} < r \leq \tau_{3,0} \\ e^{-(\tau_{3,0}-\tau_{2,0})c_2}\left\{ e^{(\tau_{2,0}-\tau_{1,0})c_1} W^{\eta}(\tau_{1,0}) + \int_{\tau_{1,0}}^{\tau_{2,0}} e^{(\tau_{2,0}-s)c_1} dW^{\eta}(s)\right\} & \\ \quad + \int_{\tau_{2,0}}^{\tau_{3,0}} e^{-(\tau_{3,0}-s)c_2} dW^{\eta}(s) + W^{\eta}(r) - W^{\eta}(\tau_{3,0}) & r > \tau_{3,0}. \end{cases}$$

□

The limiting representation given in (4) in Theorem 1 applies for the most general case of H_1 with $\delta_{1,T} > 0$ and $\delta_{2,T} > 0$. The limit when a bubble occurs without collapse (i.e. $\delta_{1,T} > 0$, $\delta_{2,T} = 0$) is readily obtained by setting $c_2 = 0$ in the above expressions, while the limit distribution under the null hypothesis, H_0, obtains by setting $c_1 = c_2 = 0$. In the homoskedastic case, where $W^{\eta}(r) = W(r)$, it can be shown that, in the limit, $\Pr(PWY > k)$, for any constant k, is an increasing function of the explosive parameter c_1, other things being equal. Essentially, this arises as a consequence of the behaviour in the second regime of $K^{\eta}_{c_1,c_2}(r)$ in $L^{\eta}_{c_1,c_2}(\tau)$, and on setting $\tau = \tau_{2,0}$. We conjecture that a similar result holds for the heteroskedastic processes covered by Assumption 1, although a proof of a result of this level of generality appears analytically intractable.

The limit distribution of PWY under the null hypothesis H_0 is given by $S^{\eta}_{0,0}$, i.e. where $K^{\eta}_{c_1,c_2}(r) = W^{\eta}(r)$. We now consider the asymptotic size of PWY for various forms of the volatility function $\omega(s)$, $s \in [0, 1]$, to assess the impact of different volatility specifications on the reliability of the test. Specifically, we consider the following cases:

Case A. Single volatility shift: $\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)1(s > \tau_{\sigma})$, $\tau_{\sigma} \in \{0.3, 0.5, 0.7\}$. *Here volatility shifts permanently from* σ_0 *to* σ_1 *at break fraction* $s = \tau_{\sigma} \in \{0.3, 0.5, 0.7\}$.

Case B. Double volatility shift: $\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)1(0.4 < s \leq 0.6)$. *Here volatility shifts temporarily from* σ_0 to σ_1 *at break fraction* $s = 0.4$ *and then reverts to* σ_0 *at break fraction* $s = 0.6$.

Case C. Logistic smooth transition in volatility: $\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)\frac{1}{1+\exp\{-50(s-0.5)\}}$. *In this case, volatility changes smoothly from* σ_0 *to* σ_1 *with a transition midpoint of* $s = 0.5$. *The speed of transition parameter (*50*) dictates that virtually all of the transition occurs between* $s = 0.4$ *and* $s = 0.6$.

Case D. Trending volatility: $\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)s$. *Here volatility follows a linear trend from* σ_0 *when* $s = 0$ *to* σ_1 *when* $s = 1$.

For each of these volatility functions we simulate the asymptotic sizes of nominal 0.05-level PWY tests, using the limit critical value obtained under homoskedasticity. We consider the range of values $\sigma_1/\sigma_0 \in \{1/6, 1/5, ..., 1/2, 1, 2, 3, ..., 6\}$, the setting $\sigma_1/\sigma_0 = 1$ giving the homoskedastic case, so the test will always have asymptotic size of 0.05 here. The sizes are computed using direct simulation of the limiting functionals appearing in Theorem 1, using 5,000 Monte Carlo replications, and approximating the Brownian motion processes in the limiting functionals using $NIID(0, 1)$ random variates, with the integrals approximated by normalized sums of 1,000 steps.

Results for the single volatility shift are given in Figure 1 (a), (c), (e). When $\sigma_1/\sigma_0 < 1$, some very modest under-sizing is observed, most evident for the earliest break fraction, $\tau_\sigma = 0.3$. It is also evident that the degree of under-sizing actually varies very little with the magnitude of σ_1/σ_0. What is far more significant, however, is the over-size present when $\sigma_1/\sigma_0 > 1$. This increases rapidly with σ_1/σ_0, pretty much irrespective of the break fraction, τ_σ, up to values of around 0.70 when $\sigma_1/\sigma_0 = 6$. Figure 2 (a) shows results for the double volatility shift. A temporary downward shift in volatility in this central region has little discernible effect on size, while an upward shift again results in very serious over-sizing, following a very similar pattern to that of the single upward shift in Figure 1. The same is true for the smooth transition in volatility shown in Figure 2 (c), which is very similar to the single instantaneous shift case of Figure 1 (c) where $\tau_\sigma = 0.5$. The trending volatility case reported in Figure 2 (e) also exhibits qualitatively similar behaviour to the shift/transition in volatility cases, and while the upward size distortions are somewhat less exaggerated than in the other cases, asymptotic size still exceeds 0.40 when $\sigma_1/\sigma_0 = 6$.

What the asymptotic size results presented above imply is that the impact of changing volatility on the size of the PWY test can be quite severe. The effect of the volatility change is very strongly dependent on the direction of the shift. If a unit root series exhibits some form of downward shift in volatility at some point in, or indeed throughout, its evolution, then spurious rejections of the unit root null are unlikely to arise. On the other hand, an upward shift can very easily lead to spurious rejections of the unit root null in favour of the explosive alternative, erroneously suggesting the presence of a bubble. The asymptotic results of Theorem 1 shed little light as to why we should find this very marked asymmetry in size behaviour between downward and upward patterns of changing volatility. Clearly, however, the combination of low volatility followed by high volatility values of Δy_t, and the forward looking nature of the DF regression in (3), is able to produce uncommonly (relative to the homoskedastic case) large positive values of $\hat{\phi}_\tau$ and DF_τ, for at least some τ, effecting the over-sizing we observe. Because we cannot realistically consider upward volatility shifts to be any less likely than downward shifts (indeed their empirical relevance would appear unquestionable), we cannot be confident that application of standard critical values for the PWY test will deliver a size-controlled procedure in the presence of non-stationary volatility. In the next section we therefore consider a wild bootstrap procedure intended to overcome these shortcomings.

5 Wild Bootstrap PWY Tests and their Asymptotic Properties

As demonstrated in the previous section, non-stationary volatility introduces a time deformation aspect to the limiting distributions of the PWY statistic which alters its form *vis-à-vis* the homoskedastic case. In this section we consider bootstrap analogues of the PWY tests based on the wild bootstrap scheme. As we shall demonstrate, this allows us to construct bootstrap unit root tests that are asymptotically robust under the null to non-stationary volatility of the form given in Assumption 1.

5.1 The Wild Bootstrap Algorithm

Our approach involves applying a wild bootstrap re-sampling scheme (see, *inter alia*, Wu, 1986; Liu, 1988; Mammen, 1993) to the first differences of the raw data and, as we will show, allows us to construct bootstrap analogues of the PWY test which are asymptotically robust to non-stationary volatility. In the context of the present problem, the wild bootstrap scheme is required rather than a standard residual re-sampling scheme, such as the i.i.d. bootstrap, because unlike these, the wild bootstrap can replicate the pattern of heteroskedasticity present in the shocks; see the discussion immediately following Algorithm 1 below.

The following steps constitute our proposed bootstrap algorithm:

Algorithm 1

Step 1. Generate T bootstrap innovations ε_t^*, as follows: $\varepsilon_1^* = 0$, $\varepsilon_t^* = w_t \Delta y_t$, $t = 2, ..., T$, where $\{w_t\}_{t=2}^T$ denotes an independent $N(0,1)$ sequence.

Step 2. Construct the bootstrap sample as the partial sum process defined by

$$y_t^* := \sum_{j=1}^{t} \varepsilon_j^*, \quad t = 1, ..., T.$$

Step 3. Compute the bootstrap test statistic

$$PWY^* := \sup_{\tau \in [\tau_0, 1]} DF^*_\tau$$

where DF^*_τ is the t-ratio on $\hat{\phi}^*_\tau$ in the fitted OLS regression

$$\Delta y^*_t = \hat{\alpha}^*_\tau + \hat{\phi}^*_\tau y^*_{t-1} + \hat{\varepsilon}^*_{t,\tau}$$

calculated over the sub-sample period $t = 1, ..., \lfloor \tau T \rfloor$, i.e.

$$DF^*_\tau = \frac{\hat{\phi}^*_\tau}{\sqrt{\hat{\sigma}^{*2}_\tau / \sum_{t=2}^{\lfloor \tau T \rfloor} \left(y^*_{t-1} - \bar{y}^*_\tau\right)^2}}$$

where $\bar{y}^*_\tau = (\lfloor \tau T \rfloor - 1)^{-1} \sum_{t=2}^{\lfloor \tau T \rfloor} y^*_{t-1}$ and $\hat{\sigma}^{*2}_\tau = (\lfloor \tau T \rfloor - 3)^{-1} \sum_{t=2}^{\lfloor \tau T \rfloor} \hat{\varepsilon}^{*2}_{t,\tau}$.

Step 4. Bootstrap p-values are computed as: $p^*_T := 1 - G^*_T(PWY)$, where $G^*_T(\cdot)$ denotes the conditional (on the original sample data) cumulative density function (cdf) of PWY^*. Notice, therefore, that the bootstrap test, run at the ξ significance level, based on PWY is then defined such that it rejects the unit root null hypothesis, H_0, if $p^*_T < \xi$. □

Notice that the bootstrap innovations ε^*_t replicate the pattern of heteroskedasticity present in the original innovations because, conditionally on Δy_t, ε^*_t is independent over time with zero mean and variance $(\Delta y_t)^2$. In practice the cdf $G^*_T(\cdot)$ required in Step 4 of Algorithm 1 will be unknown but can be approximated in the usual way through numerical simulation. This is achieved by generating N (conditionally) independent bootstrap statistics, say PWY^*_b, $b = 1, ..., N$, computed as in Algorithm 1 above. The simulated bootstrap p-value is then computed as $\tilde{p}^*_T = N^{-1} \sum_{b=1}^{N} 1\left(PWY^*_b > PWY\right)$, and is such that $\tilde{p}^*_T \overset{a.s.}{\to} p^*_T$ as $N \to \infty$. An approximate standard error for $\tilde{p}^*_T$ is given by $(\tilde{p}^*_T(1 - \tilde{p}^*_T)/N)^{1/2}$; see Hansen (1996, p.419). For a discussion on the choice of N see, *inter alia*, Davidson and MacKinnon (2000).

5.2 Asymptotic Properties

In Theorem 2, we now detail the large sample behaviour of the wild bootstrap PWY^* statistic from Algorithm 1 under both H_0 and H_1.

Theorem 2. *Under the conditions of Theorem 1, $PWY^* \overset{w}{\to}_p S^\eta_{0,0}$.* □

A comparison of the result for PWY^* in Theorem 2 with that given for PWY in Theorem 1 demonstrates the usefulness of the wild bootstrap; as the number of observations increases, the wild bootstrapped statistic has the same first-order null distribution as the original test statistic. From this result it follows, using the same arguments as in the proof of Theorem 5 in Hansen (2000), that the wild bootstrap p-values are (asymptotically) uniformly distributed under the unit root null hypothesis, leading to tests with (asymptotically) correct size in the presence of conditional heteroskedasticity of the form given in Assumption 1. The wild bootstrap procedure based on comparing PWY with bootstrap critical values (henceforth referred to as the PWY^* test) is therefore robust to non-stationary volatility in terms of asymptotic size, as seen in Figures 1-2 where the asymptotic size of PWY^* is 0.05 for all σ_1/σ_0 settings. Most importantly, this asymptotic size robustness property removes the potential for spurious rejections of H_0 in favour of a bubble when permanent upward changes in volatility occur but the series does not contain an explosive autoregressive episode.

The result in Theorem 2, taken together with the result in Theorem 1, also implies immediately that under Assumption 1 the wild bootstrap PWY^* test will also attain the same asymptotic local power function as a size-adjusted implementation of the PWY test, where the null critical values used for the latter are (infeasibly) adjusted to account for any heteroskedasticity present in the innovations. Hence we would anticipate that the finite sample power of PWY^* should be approximately the same as the size-adjusted power of PWY. In the case where volatility is constant, Theorem 2 also then implies that there is no loss in asymptotic power, relative to using the PWY test, from using its wild bootstrap analogue.

5.3 Fixed Magnitude Bubble Alternatives

Thus far we have considered alternatives H_1 where the magnitude of the bubble (and collapse) parameters are local-to-zero. Under this specification, the simple wild bootstrap procedure from Algorithm 1 based on the first differences of y_t is sufficient to allow the bootstrap test statistic to recover the null distribution associated with PWY in large samples. However, it is also important to consider the impact of non-local bubble magnitudes, and we now examine the behaviour of PWY and PWY^* under H_1 when the $\delta_{i.T}$ parameters are of a fixed magnitude, that is, $\delta_{i,T} = \delta_i > 0$, $i = 1, 2$. To that end, the large sample behaviour of the PWY and PWY^* statistics in this context is given in the following theorem.

Theorem 3. *Let $\{y_t\}$ be generated according to* (1)-(2) *under Assumption 1 and with $\delta_{i,T} = \delta_i > 0$, $i = 1, 2$. Then as $T \to \infty$, so PWY diverges to $+\infty$ at a rate as least as fast as $\lfloor \tau_{1,0} T^{1/2} \rfloor (1 + \delta_1)^{(\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor)}$, while PWY* is of $O_p(T^{1/2})$.* □

The practical consequence of the results in Theorem 3 is that the bootstrap PWY^* test is consistent against fixed alternatives. This holds because the bootstrap statistic (and, hence, the bootstrap critical values) diverge at a polynomial rate in T, whereas the original PWY statistic diverges (to $+\infty$) at an exponential rate in T. While this establishes the consistency of the PWY^* test against fixed magnitude bubble alternatives it also shows that the rate of consistency for the bootstrap test is slower than that for the original PWY test. This opens up the possibility that the bootstrap test may not be as powerful as the standard test, even where the latter is (infeasibly) adjusted to account for heteroskedasticity. However, given the significant difference between the rates of divergence of the PWY and PWY^* statistics under fixed magnitude alternatives we might expect the loss in power to be rather small in practice.

The issue raised above is a consequence of the fixed magnitude nature of the bubble (and collapse) parameters rather than the presence of non-stationary volatility in the data, since the rates of divergence of the statistics given in Theorem 3 clearly do not depend on whether the innovations are heteroskedastic or homoskedastic. To that end, assuming for now that the data were homoskedastic, an alternative bootstrap scheme that achieves the same rate of consistency and asymptotic local power function as the (infeasible) size-adjusted PWY test under fixed and local-to-zero bubble magnitude alternatives, respectively, can be constructed by utilising the BIC model selection procedure of HLS outlined in section 3. To that end, let $\hat{\varepsilon}_{Bt}$, $t = 2, ..., T$, denote the residuals from the BIC-selected estimated model using the HLS procedure. This alternative bootstrap scheme involves modifying Step 1 of Algorithm 1 to generate bootstrap innovations using $\hat{\varepsilon}_{Bt}$, $t = 2, ..., T$, in place of Δy_t, $t = 2, ..., T$. In the context of the general model (1)-(2), HLS show their procedure guarantees that the correct model and the correct regime change dates will be identified in the limit in the presence of a fixed magnitude bubble (and collapse) and, hence, that the effect of the bubble/collapse will be purged from the $\hat{\varepsilon}_{Bt}$ residuals in sufficiently large samples. As a consequence, and denoting the wild bootstrap statistic which results from this scheme by PWY_B^*, it can be seen that under the conditions of Theorem 3 and assuming homoskedasticity, $PWY_B^* \xrightarrow{w}_p S_{0,0}$, where $S_{0,0}$ is used to denote $S_{0,0}^{\eta}$ from Theorem 1 when $\omega(s) = \sigma$, for all s, in Assumption 1. Consequently, and unlike the PWY^* test, in the limit the PWY_B^* test will be consistent at the same rate as the (infeasibly) size-adjusted PWY test under fixed bubble magnitudes. The convergence result above will also hold under both H_0 and local alternatives, H_1, of the form considered in Theorems 1 and 2, such that the large sample behaviour of PWY_B^* will parallel that of PWY^* in these cases. We conjecture that allowing heteroskedasticity of the form considered in Assumption 1 would not change these large sample results (other than by replacing $S_{0,0}$ with $S_{0,0}^{\eta}$ in the convergence result above), because the large sample properties outlined above for the BIC procedure of HLS should not be altered by the presence of such heteroskedasticity.

We will return to this issue in section 6 when we evaluate the finite sample power properties of the tests, including a comparison of the PWY^* and PWY_B^* tests.

5.4 Accounting for Serial Correlation

Finally in this section we discuss how the material given thus far can be generalised to allow for serial correlation in ε_t. To do so we allow for a very general pattern of possible weak dependence in ε_t through the linear process

$$\varepsilon_t = C(L)\sigma_t z_t = \sum_{j=0}^{\infty} C_j \sigma_{t-j} z_{t-j}$$

where $C(z)$ is assumed to satisfy standard summability and invertibility conditions, *viz.*, $\sum_{j=0}^{\infty} j|C_j| < \infty$ and $C(z) \neq 0$ for all $|z| \leq 1$, respectively. These conditions are satisfied, for example, by all stable and invertible finite-order ARMA processes. In this case, provided that the sub-sample regressions (3) used to construct PWY are augmented by inclusion of the lagged-difference regressors $\Delta y_{t-1}, ..., \Delta y_{t-p}$, where p is chosen such that, as $T \to \infty$, $1/p + p^3/T \to 0$, it can be shown that the asymptotic results regarding PWY in Theorem 1 remain unaltered. Moreover, none of the large sample results stated in this section relating to the bootstrap PWY^* and PWY_B^* statistics are reliant on the absence of serial correlation in ε_t; this is the case because the wild bootstrap re-sampling device used in Step 1 of Algorithm 1 annihilates any weak dependence present in either Δy_t in the context of PWY^* or $\hat{\varepsilon}_{Bt}$ in the context of PWY_B^*. A particular implication of this is that there is no requirement to augment the sub-sample regressions underlying the bootstrap procedures PWY^* and PWY_B^* with lagged-difference regressors for the foregoing large sample properties of these two bootstrap procedures to continue to hold when ε_t is weakly dependent. In fact, the results stated in Theorems 2 and 3 remain valid for any lag length, p^* say, in the bootstrap analogue of (3) such that $p^*/T^{1/3} \to 0$ as $T \to \infty$. Re-coloured versions of our bootstrap procedures, constructed along the lines considered in Cavaliere and Taylor (2009a), could therefore also be considered.

6 Finite Sample Properties

In this section we use Monte Carlo simulation methods to compare the finite sample size and power properties of the original PWY test and the two new bootstrap tests PWY^* and PWY_B^* proposed in section 5 (all constructed without lagged-difference augmentation). Using the DGP (1)-(2) we set $\mu = 0$ (without loss of generality), $u_1 = \varepsilon_1$ and generate z_t as $IIDN(0, 1)$, for a sample of size $T = 200$. Here $\varepsilon_t = \sigma_t z_t$, where $\sigma_t = \omega(t/T)$ is the discrete time analogue of the volatility functions given by Cases A-D in section 4. All simulations are conducted at the nominal asymptotic 0.05 level using 5,000 Monte Carlo replications and $N = 499$ bootstrap replications. In sections 6.1 and 6.2 below, we present results for finite sample size and power, respectively.

6.1 Empirical Size

Finite sample (empirical) size results for PWY, PWY^* and PWY_B^* are given in Figure 1 (b), (d), (f) and Figure 2 (b), (d), (f). When $\sigma_1/\sigma_0 < 1$, they each bear a close resemblance to their asymptotic counterparts in Figure 1 (a), (c), (e) and Figure 2 (a), (c), (e), respectively. When $\sigma_1/\sigma_0 > 1$, on the whole the same is also true, although PWY^* and PWY_B^* are now both a little over-sized, with PWY_B^* generally being the more distorted of the two. Of course, it is when $\sigma_1/\sigma_0 > 1$ that the bootstrap tests have their work cut out in attempting to mimic a size-correction exercise for PWY, since the latter has greatly inflated size in these situations. Taking into account the magnitude of the over-size present for PWY, it is fair to say that both PWY^* and PWY_B^* are doing a very effective job in terms of controlling finite sample size here. Overall, it is encouraging to see that the predictions of the asymptotic theory provided in section 5 for the proposed bootstrap test procedures broadly carry over to sample sizes of empirical relevance. As such, the wild bootstrap PWY^* and PWY_B^* tests provide an approach to testing for the presence of a bubble that, unlike the original PWY test, are not susceptible to spuriously indicating the presence of a bubble when the series follows a unit root process driven by innovations which display some form of increasing volatility.

6.2 Empirical Power

To examine the finite sample powers of PWY, PWY^* and PWY_B^*, we consider the set of bubble magnitudes $\delta_{1,T} = \delta_1 \in \{0.02, 0.04, 0.06, 0.08\}$ for the non-collapsing case $\delta_{2,T} = 0$, along with the bubble regime timings $\tau_{1,0} = 0.4$ and $\tau_{2,0} = 0.6$. This setting, combined with the volatility functions given in Cases B or C of section 4, represent examples where the volatility changes occur at (or around) the start and end of the bubble regime, while for the volatility functions given in Cases A and D, the volatility change timings are unrelated to the timing of the bubble. In cases where a simulated DGP resulted in a downward explosive regime (i.e. if $y_{\lfloor \tau_{2,0}T \rfloor} < y_{\lfloor \tau_{1,0}T \rfloor}$), typically due to the explosive period originating with negative values, we multiplied the simulated series by -1, so as to ensure that all generated series had upward explosive regimes. Tables 1 and 2 report results for the discrete time analogues of the four volatility functions in Cases A-D for the settings $\sigma_1/\sigma_0 \in \{1/6, 1/3, 1, 3, 6\}$; the case $\delta_1 = 0$ is also included to represent size. All three tests have different finite sample sizes, depending on the volatility function A-D. While, when comparing power properties, the differences between the empirical sizes of the two bootstrap procedures PWY^* and PWY_B^* can largely be ignored, the differences between their empirical sizes and those of the original PWY test cannot, since the potential for PWY to be over-sized would render raw power comparisons meaningless. In what follows then, in addition to reporting the raw powers of PWY, we also report two additional (infeasible) size-adjusted powers for PWY. These are PWY_1^{adj} and PWY_2^{adj} which, for a given volatility function, are the powers of PWY when its size is adjusted to match those of PWY^* and PWY_B^*, respectively.

In Table 1 we provide finite sample (empirical) powers under the discrete time analogue of volatility function A, the single volatility shift. The first thing to note is that, across all the volatility settings, each of the tests has power that rises monotonically with δ_1. The powers of PWY^* and PWY_1^{adj} are always very close to each other, as are those of PWY_B^* and PWY_2^{adj}. Essentially then, neither PWY^* nor PWY_B^* lose any finite sample power relative to the appropriately size-adjusted PWY test. It is also important to note that the powers of PWY^* and PWY_B^* are always very similar. In those few cases where PWY_B^* appears a little more powerful than PWY^*, generally for the smaller values of δ_1, this could quite reasonably be ascribed to the former's slightly higher corresponding empirical size. It is also worth noting that the timing of the volatility shift appears to have little effect on power. On the other hand (with the exception of the unadjusted PWY test), upward volatility shifts do appear to be associated with lower levels of power, relative to the homoskedastic case or the downward volatility shift cases. Of course, these represent cases where the size of the original PWY test is very high, so this is perhaps hardly surprising. Table 2 shows the results for the volatility functions B-D. In each case we again see that the empirical powers of PWY^*, PWY_B^*, PWY_1^{adj} and PWY_2^{adj} increase with δ_1. As was the case with volatility function A, the powers are all fairly similar among these tests, and upward shifts in volatility, which cause PWY to be over-sized, are associated with lower levels of power. It appears that the specific form of volatility only affects the power of PWY^* and PWY_B^* in as much as it affects the size of PWY. That we observe the powers of PWY^* and PWY_B^* to be so similar also suggests that, in practice, the simpler PWY^* procedure gives away little or nothing to its more elaborate model-based counterpart, and that the potential power issue under a fixed bubble magnitude specification discussed in section 5.3 does not appear to be a concern in practice.

Taking our finite sample size and power results together, we find that both of our proposed wild bootstrap procedures PWY^* and PWY_B^* are effective in restoring size control in the presence of non-stationary volatility whilst simultaneously maintaining available levels of power. While PWY_B^* has certain theoretical power advantages over PWY^*, this property does not appear to translate into any discernible finite sample power advantage in practice. An argument could also be made for using PWY^* on the grounds that it is not model-based and does not require a particular specification for the collapse regime, nor does it require a unit root regime prior to the onset of a bubble (i.e. the bubble regime can begin at the very start of the process), unlike PWY_B^*. However, even if PWY_B^* mis-specifies the bubble regime in some way, the fact that PWY^* has the same power levels as PWY_B^* suggests that precise modelling of the bubble and collapse regimes is not critical for competitive power, so potential concerns regarding, say, the nature of a bubble's collapse, are not pertinent in this problem. Overall, given the very close similarity in performance between PWY^* and PWY_B^*, we would recommend the use of PWY^* in practice, since it is simpler to compute, and has marginally better finite sample size properties.

7 An Empirical Illustration

As an empirical application of our new bootstrap approach we consider several commodity price time series. The demand for many primary and intermediate commodities increased substantially between the end of the dot-com crash in 2001 and the 2007-2009 global financial crisis, driven by strong global economic growth over this period (with particularly strong growth in the BRIC countries). As a consequence, for many commodities over this period significant price rises occurred, followed by significant price falls as a consequence of the 2007-2009 financial crisis. This feature of commodity prices has led several researchers to employ PWY-type tests to investigate the possibility that some commodity price series over this period may have contained periods of explosive autoregressive behaviour consistent with the presence of a speculative bubble.[5] Applying the PWY test to data from 2000-2009, Gilbert (2010) finds strong evidence of bubbles in the copper market, weaker evidence of bubbles in the nickel and crude oil market, and no evidence of bubbles in the aluminium market. Using a modified version of the PWY test, Phillips and Yu (2010, 2011) find evidence of explosive autoregressive behaviour in the crude oil market and platinum market in 2008. Homm and Breitung (2012) apply the PWY test to two commodity price series - crude oil and gold over the period 1985-2010. They find statistically significant evidence of explosive autoregressive behaviour in the gold price series in the late-2000s, but no significant evidence for the crude oil price series. Figuerola-Ferretti *et al.* (2015) apply the Phillips *et al.* (2014) procedure to cash and futures prices of non-ferrous metals, finding evidence for multiple periods of explosive behaviour during the 2000s in five out of the six metals considered.

Many of the papers which test for speculative bubbles in commodity prices using PWY-type tests do so using samples of data that span periods of global financial and macroeconomic instability. For example, the applied studies referred to in the previous paragraph employ samples of data that span the period before the 2007-2009 financial crisis when global economic growth was strong and financial market volatility was generally quite low, and the period during and immediately after the financial crisis when there was a high level of uncertainty in financial markets and when many countries experienced the start of a significant recession. Therefore it seems highly likely that, even if a bubble did not exist during the sample periods examined in these studies, the unconditional volatility of the first differenced price series would not be constant over the samples considered. Hence, our bootstrap approach might prove to be useful in this context.

A further reason to advocate the use of our bootstrap approach when testing for bubbles in commodity prices is that several previous empirical studies of commodity price volatility over this period have indeed found statistically significant evidence of non-constancy in the unconditional volatility of the first differenced series. For example, using long-run monthly data that includes the 2007-2009 crisis period Calvo-Gonzalez *et al.* (2010) employ the CUSUM methodology of Inclan and Tiao (1994) and Kokoszka and Leipus (1999) to search for structural breaks in the unconditional volatility of the returns series for 45 commodities. They find statistically significant evidence of breaks for many of the commodities examined. Interestingly, Calvo-Gonzalez *et al.* (2010) conclude that the timing of the structural breaks detected is idiosyncratic and that there is no consistent pattern to the results across the individual commodities. Also using CUSUM-type tests, Ewing and Malik (2010) find evidence of multiple volatility breaks in daily data on the WTI crude oil price over the period 1993-2008, including a break at the start of the 2007-2009 financial crisis. Vivian and Wohar (2012) apply CUSUM-type tests to investigate structural breaks in the volatility of daily returns for 28 commodities

[5] Note that for commodity prices the underlying fundamental (equivalent to the dividend for stocks) is an unobserved 'convenience yield'. Hence studies of speculative bubbles in commodity prices typically focus only on detecting explosive autoregressive behaviour in the price series and/or the natural logarithm of the price series.

over the period 1985-2010. They find some evidence of breaks corresponding to the 2007-2009 financial crisis, but also find that structural breaks in volatility often occur in non-crisis periods. Similarly to Calvo-Gonzalez *et al.* (2010), Vivan and Wohar (2012) conclude that there is no consistent pattern to the timing of breaks across individual commodities.

The application reported here focuses on the prices of two types of crude oil (Brent and West Texas Intermediate (WTI)), three precious metals (gold, silver and platinum) and two non-ferrous metals (aluminium and copper). Results are reported for nominal weekly data and real monthly data, using the US CPI as a deflator. The oil prices are spot prices from the Energy Information Administration. The precious metals prices are spot prices from the London Bullion Market and the London Platinum and Palladium Market; the non-ferrous metal prices are three-month futures prices from the London Metals Exchange. In all cases the sample period starts at the beginning of January 2000 and finishes at the end of December 2013, giving 168 monthly observations and 731 weekly observations.[6] All of the commodity price series were downloaded using Thomson Reuters Datastream, and the CPI data was downloaded from the Federal Reserve Bank of Louis FRED database. The levels and first-differences of the series are plotted in Figures 3 and 4, for the monthly and weekly series respectively. A simple visual analysis of these plots suggests that the assumption of stationary unconditional volatility is unrealistic for these series, with commodity volatility appearing to increase over the sample period in most cases.

To investigate more formally for the possible presence of non-stationary volatility in these series, in Table 3 we report results from application of the stationary volatility tests of Cavaliere and Taylor (2008b, pp. 311–312). We apply all four of their proposed tests ($\mathcal{H}_{KS}$, $\mathcal{H}_R$, $\mathcal{H}_{CVM}$ and $\mathcal{H}_{AD}$, using a Bartlett long run variance estimator with lag truncation parameter 4), and to mitigate possible confounding effects of any bubble/collapse that might be present, for each series we compute the tests employing the fitted residuals from the BIC-selected bubble model of HLS outlined in section 3, i.e. $\hat{\varepsilon}_{Bt}$ of section 5.3. It can be seen that for each of the commodities and for both the monthly and weekly series there is statistically significant evidence against the null of stationary volatility from at least one of the tests at conventional significance levels. As might be expected given the relative sample sizes involved, the evidence is stronger for the weekly data than for the monthly data, with a rejection delivered by all of the tests at the 0.01-level when the former is used. For the monthly data the $\mathcal{H}_{AD}$ test yields the most evidence among the four tests for non-stationary volatility in the data, while the $\mathcal{H}_R$ test provides the least. Interestingly, the Monte Carlo simulations reported in Cavaliere and Taylor (2008b) reveal that when there is a single discrete break in volatility, or when volatility follows a linear trend, the $\mathcal{H}_{AD}$ test has the greatest finite sample power and is noticeably more powerful than the $\mathcal{H}_R$ test, which is the least powerful of the four tests. In contrast for the case

[6] Note that in some of the previous research on commodity price bubbles using PWY-type tests raw prices are used, whilst in other studies the natural logarithms of the prices are used. Here we report results using raw price data.

of two discrete structural breaks in volatility, Cavaliere and Taylor (2008b) find that the $\mathcal{H}_R$ test is now the most powerful of the four tests. Thus, when considered alongside the simulation results in Cavaliere and Taylor (2008b), the overall pattern of the results in Table 3 suggests that perhaps a single discrete break volatility model, or a trending volatility model, is more likely to be appropriate than a multiple break volatility model.

We now turn to testing for the presence of speculative bubbles in the commodities data. To that end, for each series we report in Table 4 the outcome of the PWY test statistic, along with the corresponding p-value (calculated under the assumption of homoskedastic errors by Monte Carlo simulation) for the standard PWY test, together with wild bootstrap p-values, computed as in Algorithm 1 and also using the BIC-based variant of Algorithm 1 discussed in section 5.3, using $N = 9999$ bootstrap replications in each case. In computing the PWY statistics we allowed for a maximum of six lagged-differenced regressors to account for serial correlation and selected the lag length by BIC (the sub-sample regressions underlying the bootstrap procedures PWY^* and PWY^*_B do not include lagged-difference augmentation; see the discussion in section 5.4).

The results in Table 4 show that the standard PWY test rejects the unit root null hypothesis in favour of the explosive alternative at the 0.01-level for all of the series considered in both the monthly and weekly data, in each case providing very strong evidence for the presence of a speculative bubble; indeed, most of the p-values are at or very close to zero. When considering the wild bootstrap PWY^* and PWY^*_B tests, which are robust to the presence of non-stationary volatility in the data, we find much less emphatic evidence for speculative bubble behaviour. For the monthly data, 0.01-level rejections are obtained by the PWY^* and PWY^*_B tests only for the copper series. Moreover, the PWY^* test (which displayed the best finite sample size control in the simulations) does not deliver rejections at the 0.05-level for any other commodity, although some weaker 0.10-level rejections are obtained for Brent oil, gold, silver and aluminium. The PWY^*_B p-values are always lower than the corresponding PWY^* p-values, as might be expected given their finite sample properties, and more evidence in favour of a bubble is therefore found by this test at the 0.05- and 0.10-levels. For the case of weekly data, no 0.01-level rejections are found by PWY^*, while a 0.01-level rejection is only obtained by PWY^*_B for copper. At the 0.05-level, the PWY^* test rejects for Brent oil, silver and copper, while evidence for a bubble is only found at the 0.10-level for the remaining series. Again, the PWY^*_B p-values are found to be lower than those for PWY^*, but no additional 0.05-level rejections are seen.

Our results therefore show that the use of critical values that are robust to the presence of non-stationary unconditional volatility lead to much less clear evidence of bubble behaviour than when critical values are used that assume stationary unconditional volatility. The fact that the bootstrap PWY^* and PWY^*_B tests control size under non-stationary volatility yet do not lose power relative to the PWY tests under stationary volatility (see the Monte Carlo results reported in sections 6.2 and 6.3), combined with the results from Table 3 where the hypothesis of stationary volatility is rejected for all of the series, would suggest that the standard PWY

test results are likely to be an overstatement of the evidence for a bubble in these series, with the more equivocal findings of the PWY^* and PWY^*_B tests providing a more reliable indicator of the presence or absence of speculative bubbles. Finally, note that the tests we have considered are designed for detection of a single bubble episode, although the possibility exists that multiple explosive periods could be present in some of these series, as considered by Figuerola-Ferretti *et al.* (2015) under an implicit assumption of homoskedasticity.

8 Conclusions

In this paper we have explored the impact that non-stationary volatility has on the performance of the test for explosive financial bubbles based on sub-sample Dickey-Fuller statistics proposed in Phillips *et al.* (2011). Numerical and analytical evidence was presented that showed that empirically relevant models of non-stationary volatility can have potentially serious implications for the reliability of this test, with size often being substantially above the nominal level, thereby giving rise to spurious indications of explosive behaviour in the data. To address this problem we have considered wild bootstrap-based implementations of the Phillips *et al.* (2011) test, these having proved to be highly successful in left-tailed unit root testing applications. Our bootstrap tests have the considerable advantage that they are not tied to a given parametric model of volatility within the class of non-stationary volatility processes considered. The asymptotic validity of our bootstrap tests within the class of non-stationary volatility considered was demonstrated and Monte Carlo simulation evidence was provided which showed them to be effective in controlling finite sample size under non-stationary volatility. Moreover, the bootstrap tests were found not to sacrifice power relative to infeasibly size-correcting the original test. We also provided an empirical application involving commodity price time series and found considerably less clear evidence for the presence of speculative bubbles in these data when using our wild bootstrap implementations of the Phillips *et al.* (2011) test.

It is important to note that while our paper has focused on testing for the presence of a bubble using the Phillips *et al.* (2011) test, the subsequent dating procedure proposed by these authors also relies on recursive Dickey-Fuller statistics, the distributions of which will be similarly affected by non-stationary volatility. As a result, the dating properties of the procedure will also be sensitive to departures from homoskedasticity. We envisage that a wild bootstrap correction could be gainfully employed here also. Finally, the testing procedure developed in Phillips *et al.* (2011) is directed towards processes which, under the alternative hypothesis, admit one explosive sub-sample regime. In a more recent paper, Phillips *et al.* (2014) extends the methodology of the Phillips *et al.* (2011) to allow for the possibility of multiple bubble regimes under the alternative. This approach involves a test for the presence of at least one bubble based on a double supremum of forwards and backwards recursive Dickey-Fuller statistics, followed by subsequent consideration of the sequence of supremum backward recursive Dickey-Fuller statistics to both determine the number of bubbles present and date stamp their timings. Since both the double supremum test statistic and the subsequent bubble identification procedure are based on recursive Dickey-Fuller statistics, once again the challenges raised in this paper regarding the impact of non-stationary volatility on inference will also be germane. We would fully expect, however, that the wild bootstrap methodology outlined in this paper could also be demonstrated to deliver a robust approach to bubble detection and dating in this more general setting.

A Appendix

Without loss of generality we can set $\mu = 0$ and $u_1 = 0$ in what follows.

A.1 Proof of Theorem 1

By backward substitution in (1) we obtain

$$y_t = \begin{cases} \sum_{i=1}^{t} \varepsilon_i & t = 2, ..., \lfloor \tau_{1,0} T \rfloor \\ (1+\delta_{1,T})^{t - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{1,0} T \rfloor + 1, ..., \lfloor \tau_{2,0} T \rfloor \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{t} (1+\delta_{1,T})^{t-i} \varepsilon_i & \\ (1-\delta_{2,T})^{t - \lfloor \tau_{2,0} T \rfloor} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{2,0} T \rfloor + 1, ..., \lfloor \tau_{3,0} T \rfloor \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i \} & \\ \quad + \sum_{i=\lfloor \tau_{2,0} T \rfloor + 1}^{t} (1-\delta_{2,T})^{t-i} \varepsilon_i & \\ (1-\delta_{2,T})^{\lfloor \tau_{3,0} T \rfloor - \lfloor \tau_{2,0} T \rfloor} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{3,0} T \rfloor + 1, ..., T \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i \} & \\ \quad + \sum_{i=\lfloor \tau_{2,0} T \rfloor + 1}^{\lfloor \tau_{3,0} T \rfloor} (1-\delta_{2,T})^{\lfloor \tau_{3,0} T \rfloor - i} \varepsilon_i + \sum_{i=\lfloor \tau_{3,0} T \rfloor + 1}^{t} \varepsilon_i & \end{cases} \tag{A.1}$$

and subsequently

$$T^{-1/2} y_{\lfloor rT \rfloor} =$$

$$T^{-1/2} \begin{cases} \sum_{i=1}^{\lfloor rT \rfloor} \varepsilon_i & \lfloor rT \rfloor = 2, ..., \lfloor \tau_{1,0} T \rfloor \\ (1+\delta_{1,T})^{\lfloor rT \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & \lfloor rT \rfloor = \lfloor \tau_{1,0} T \rfloor + 1, ..., \lfloor \tau_{2,0} T \rfloor \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor rT \rfloor} (1+\delta_{1,T})^{\lfloor rT \rfloor - i} \varepsilon_i & \\ (1-\delta_{2,T})^{\lfloor rT \rfloor - \lfloor \tau_{2,0} T \rfloor} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & \lfloor rT \rfloor = \lfloor \tau_{2,0} T \rfloor + 1, ..., \lfloor \tau_{3,0} T \rfloor \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i \} & \\ \quad + \sum_{i=\lfloor \tau_{2,0} T \rfloor + 1}^{\lfloor rT \rfloor} (1-\delta_{2,T})^{\lfloor rT \rfloor - i} \varepsilon_i & \\ (1-\delta_{2,T})^{\lfloor \tau_{3,0} T \rfloor - \lfloor \tau_{2,0} T \rfloor} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & \lfloor rT \rfloor = \lfloor \tau_{3,0} T \rfloor + 1, ..., T. \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i \} & \\ \quad + \sum_{i=\lfloor \tau_{2,0} T \rfloor + 1}^{\lfloor \tau_{3,0} T \rfloor} (1-\delta_{2,T})^{\lfloor \tau_{3,0} T \rfloor - i} \varepsilon_i + \sum_{i=\lfloor \tau_{3,0} T \rfloor + 1}^{\lfloor rT \rfloor} \varepsilon_i & \end{cases}$$

Under $\delta_i = c_i/T$, for $0 < a < b < 1$, $(1+\delta_1)^{\lfloor bT \rfloor - \lfloor aT \rfloor} = e^{(b-a)c_1} + o(1)$ and $(1-\delta_2)^{\lfloor bT \rfloor - \lfloor aT \rfloor} = e^{-(b-a)c_2} + o(1)$, and then, following Phillips (1987) we find

$$\begin{aligned} T^{-1/2} y_{\lfloor rT \rfloor} &\xrightarrow{w} \overline{\omega} \begin{cases} W^\eta(r) & r \le \tau_{1,0} \\ e^{(r-\tau_{1,0})c_1} W^\eta(\tau_{1,0}) + \int_{\tau_{1,0}}^r e^{(r-s)c_1} dW^\eta(s) & \tau_{1,0} < r \le \tau_{2,0} \\ e^{-(r-\tau_{2,0})c_2} \left\{ e^{(\tau_{2,0}-\tau_{1,0})c_1} W^\eta(\tau_{1,0}) + \int_{\tau_{1,0}}^{\tau_{2,0}} e^{(\tau_{2,0}-s)c_1} dW^\eta(s) \right\} & \tau_{2,0} < r \le \tau_{3,0} \\ \quad + \int_{\tau_{2,0}}^r e^{-(r-s)c_2} dW^\eta(s) & \\ e^{-(\tau_{3,0}-\tau_{2,0})c_2} \left\{ e^{(\tau_{2,0}-\tau_{1,0})c_1} W^\eta(\tau_{1,0}) + \int_{\tau_{1,0}}^{\tau_{2,0}} e^{(\tau_{2,0}-s)c_1} dW^\eta(s) \right\} & r > \tau_{3,0} \\ \quad + \int_{\tau_{2,0}}^{\tau_{3,0}} e^{-(\tau_{3,0}-s)c_2} dW^\eta(s) + W^\eta(r) - W^\eta(\tau_{3,0}) & \end{cases} \\ &= \overline{\omega} K^\eta_{c_1,c_2}(r) \end{aligned}$$

Also,

$$\Delta y_t = \begin{cases} \varepsilon_t & t = 2, ..., \lfloor \tau_{1,0} T \rfloor \\ \varepsilon_{\lfloor \tau_{1,0} T \rfloor + 1} + \delta_{1,T} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{1,0} T \rfloor + 1 \\ \varepsilon_t + \delta_{1,T}(1+\delta_{1,T})^{t - \lfloor \tau_{1,0} T \rfloor - 1} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{1,0} T + 2 \rfloor, ..., \lfloor \tau_{2,0} T \rfloor \\ \quad + \delta_{1,T} \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{t-1} (1+\delta_{1,T})^{t-1-i} \varepsilon_i & \\ \varepsilon_{\lfloor \tau_{2,0} T \rfloor + 1} - \delta_{2,T}(1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{2,0} T \rfloor + 1 \\ \quad - \delta_{2,T} \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i & \\ \varepsilon_t - \delta_{2,T}(1-\delta_{2,T})^{t - \lfloor \tau_{2,0} T \rfloor - 1} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - \lfloor \tau_{1,0} T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0} T \rfloor} \varepsilon_i & t = \lfloor \tau_{2,0} T + 2 \rfloor, ..., \lfloor \tau_{3,0} T \rfloor \\ \quad + \sum_{i=\lfloor \tau_{1,0} T \rfloor + 1}^{\lfloor \tau_{2,0} T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0} T \rfloor - i} \varepsilon_i \} & \\ \quad - \delta_{2,T} \sum_{i=\lfloor \tau_{2,0} T \rfloor + 1}^{t-1} (1-\delta_{2,T})^{t-i-1} \varepsilon_i & \\ \varepsilon_t & t = \lfloor \tau_{3,0} T \rfloor + 1, ..., T \end{cases}$$

so

$$\Delta y_t = \begin{cases} \varepsilon_t & t = 2, ..., \lfloor \tau_{1,0} T \rfloor \\ \varepsilon_{\lfloor \tau_{1,0} T \rfloor + 1} + O_p(T^{-1/2}) & t = \lfloor \tau_{1,0} T \rfloor + 1 \\ \varepsilon_t + O_p(T^{-1/2}) & t = \lfloor \tau_{1,0} T + 2 \rfloor, ..., \lfloor \tau_{2,0} T \rfloor \\ \varepsilon_{\lfloor \tau_{2,0} T \rfloor + 1} + O_p(T^{-1/2}) & t = \lfloor \tau_{2,0} T \rfloor + 1 \\ \varepsilon_t + O_p(T^{-1/2}) & t = \lfloor \tau_{2,0} T + 2 \rfloor, ..., \lfloor \tau_{3,0} T \rfloor \\ \varepsilon_t & t = \lfloor \tau_{3,0} T \rfloor + 1, ..., T. \end{cases}$$

Then, since it is easily shown that $\hat{\sigma}_\tau^2 = \lfloor \tau T \rfloor^{-1} \sum_{t=1}^{\lfloor \tau T \rfloor} \Delta y_t^2 + o_p(1)$ and $\lfloor \tau T \rfloor^{-1} \sum_{t=1}^{\lfloor \tau T \rfloor} \Delta y_t^2 \xrightarrow{p} \tau^{-1} \int_0^\tau \omega(h)^2 dh$, we find that $\hat{\sigma}_\tau^2 \xrightarrow{p} \tau^{-1} \int_0^\tau \omega(h)^2 dh$. The stated limit for PWY then follows from an application of the Continuous Mapping Theorem (CMT).

A.2 Proof of Theorem 2

For brevity, we will only present results for DGP 4 under H_1. Results for DGPs 1-3 are simply obtained as special cases. According to our bootstrap algorithm,

$$T^{-1/2} y^*_{\lfloor rT \rfloor} = T^{-1/2} \sum_{j=1}^{\lfloor rT \rfloor} \varepsilon_j^*$$

with $\varepsilon_j^* = w_j \Delta y_j$. Now,

$$
w_j \Delta y_j =
\begin{cases}
w_j \varepsilon_j & j = 2, ..., \lfloor \tau_{1,0} T \rfloor \\
w_{\lfloor \tau_{1,0}T \rfloor+1} \varepsilon_{\lfloor \tau_{1,0}T \rfloor+1} + w_{\lfloor \tau_{1,0}T \rfloor+1} \delta_{1,T} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i & j = \lfloor \tau_{1,0}T \rfloor + 1 \\
w_j \varepsilon_j + w_j \delta_{1,T} (1+\delta_{1,T})^{j-\lfloor \tau_{1,0}T \rfloor - 1} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i & j = \lfloor \tau_{1,0}T + 2 \rfloor, ..., \lfloor \tau_{2,0}T \rfloor \\
+ w_j \delta_{1,T} \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{j-1} (1+\delta_{1,T})^{j-1-i} \varepsilon_i & \\
w_{\lfloor \tau_{2,0}T \rfloor+1} \varepsilon_{\lfloor \tau_{2,0}T \rfloor+1} - w_{\lfloor \tau_{2,0}T \rfloor+1} \delta_{2,T} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i & j = \lfloor \tau_{2,0}T \rfloor + 1 \\
- w_{\lfloor \tau_{2,0}T \rfloor+1} \delta_{2,T} \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - i} \varepsilon_i & \\
w_j \varepsilon_j - w_j \delta_{2,T} (1-\delta_{2,T})^{j-\lfloor \tau_{2,0}T \rfloor - 1} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i & j = \lfloor \tau_{2,0}T + 2 \rfloor, ..., \lfloor \tau_{3,0}T \rfloor \\
+ \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - i} \varepsilon_i \} & \\
- w_j \delta_{2,T} \sum_{i=\lfloor \tau_{2,0}T \rfloor+1}^{j-1} (1-\delta_{2,T})^{j-i-1} \varepsilon_i & \\
w_j \varepsilon_j & j = \lfloor \tau_{3,0}T \rfloor + 1, ..., T
\end{cases}
$$

so

$$
T^{-1/2} y^*_{\lfloor rT \rfloor} = T^{-1/2} \sum_{j=2}^{\lfloor rT \rfloor} w_j \varepsilon_j +
\begin{cases}
p_1 & \lfloor rT \rfloor = \lfloor \tau_{1,0}T \rfloor + 1 \\
p_1 + p_2 & \lfloor rT \rfloor = \lfloor \tau_{1,0}T + 2 \rfloor, ..., \lfloor \tau_{2,0}T \rfloor \\
p_1 + p_2 + p_3 & \lfloor rT \rfloor = \lfloor \tau_{2,0}T \rfloor + 1 \\
p_1 + p_2 + p_3 + p_4 & \lfloor rT \rfloor = \lfloor \tau_{2,0}T + 2 \rfloor, ..., T
\end{cases}
$$

where

$$
\begin{aligned}
p_1 &= T^{-1/2} w_{\lfloor \tau_{1,0}T \rfloor+1} \delta_{1,T} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i \\
p_2 &= T^{-1/2} \sum_{j=\lfloor \tau_{1,0}T+2 \rfloor}^{\lfloor rT \rfloor} w_j \delta_{1,T} (1+\delta_{1,T})^{j-\lfloor \tau_{1,0}T \rfloor - 1} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i \\
&\quad + T^{-1/2} \sum_{j=\lfloor \tau_{1,0}T+2 \rfloor}^{\lfloor rT \rfloor} w_j \delta_{1,T} \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{j-1} (1+\delta_{1,T})^{j-1-i} \varepsilon_i \\
p_3 &= -T^{-1/2} w_{\lfloor \tau_{2,0}T \rfloor+1} \delta_{2,T} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i \\
&\quad - T^{-1/2} w_{\lfloor \tau_{2,0}T \rfloor+1} \delta_{2,T} \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - i} \varepsilon_i \\
p_4 &= -T^{-1/2} \sum_{j=\lfloor \tau_{2,0}T+2 \rfloor}^{\lfloor rT \rfloor} w_j \delta_{2,T} (1-\delta_{2,T})^{j-\lfloor \tau_{2,0}T \rfloor - 1} \{ (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor} \sum_{i=1}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_i \\
&\quad + \sum_{i=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} (1+\delta_{1,T})^{\lfloor \tau_{2,0}T \rfloor - i} \varepsilon_i \} - T^{-1/2} \sum_{j=\lfloor \tau_{2,0}T+2 \rfloor}^{\lfloor rT \rfloor} w_j \delta_{2,T} \sum_{i=\lfloor \tau_{2,0}T \rfloor+1}^{j-1} (1-\delta_{2,T})^{j-i-1} \varepsilon_i.
\end{aligned}
$$

Also,

$$\begin{aligned}
p_1 &= T^{-1}c_1 w_{\lfloor\tau_{1,0}T\rfloor+1}T^{-1/2}\textstyle\sum_{i=1}^{\lfloor\tau_{1,0}T\rfloor}\varepsilon_i \\
&= O_p(T^{-1}) \\
p_2 &= T^{-1/2}c_1 T^{-1/2}\textstyle\sum_{j=\lfloor\tau_{1,0}T+2\rfloor}^{\lfloor rT\rfloor} w_j(1+\delta_{1,T})^{j-\lfloor\tau_{1,0}T\rfloor-1}T^{-1/2}\textstyle\sum_{i=1}^{\lfloor\tau_{1,0}T\rfloor}\varepsilon_i \\
&\quad +T^{-1/2}c_1 T^{-1/2}\textstyle\sum_{j=\lfloor\tau_{1,0}T+2\rfloor}^{\lfloor rT\rfloor} w_j T^{-1/2}\textstyle\sum_{i=\lfloor\tau_{1,0}T\rfloor+1}^{j-1}(1+\delta_{1,T})^{j-1-i}\varepsilon_i \\
&= O_p(T^{-1/2}) \\
p_3 &= -T^{-1}c_2 w_{\lfloor\tau_{2,0}T\rfloor+1}(1+\delta_{1,T})^{\lfloor\tau_{2,0}T\rfloor-\lfloor\tau_{1,0}T\rfloor}T^{-1/2}\textstyle\sum_{i=1}^{\lfloor\tau_{1,0}T\rfloor}\varepsilon_i \\
&\quad -T^{-1}c_2 w_{\lfloor\tau_{2,0}T\rfloor+1}T^{-1/2}\textstyle\sum_{i=\lfloor\tau_{1,0}T\rfloor+1}^{\lfloor\tau_{2,0}T\rfloor}(1+\delta_{1,T})^{\lfloor\tau_{2,0}T\rfloor-i}\varepsilon_i \\
&= O_p(T^{-1}) \\
p_4 &= -T^{-1/2}c_2 T^{-1/2}\textstyle\sum_{j=\lfloor\tau_{2,0}T+2\rfloor}^{\lfloor rT\rfloor} w_j(1-\delta_{2,T})^{j-\lfloor\tau_{2,0}T\rfloor-1}\{(1+\delta_{1,T})^{\lfloor\tau_{2,0}T\rfloor-\lfloor\tau_{1,0}T\rfloor}T^{-1/2}\textstyle\sum_{i=1}^{\lfloor\tau_{1,0}T\rfloor}\varepsilon_i \\
&\quad +T^{-1/2}\textstyle\sum_{i=\lfloor\tau_{1,0}T\rfloor+1}^{\lfloor\tau_{2,0}T\rfloor}(1+\delta_{1,T})^{\lfloor\tau_{2,0}T\rfloor-i}\varepsilon_i\} \\
&\quad -T^{-1/2}c_2 T^{-1/2}\textstyle\sum_{j=\lfloor\tau_{2,0}T+2\rfloor}^{\lfloor rT\rfloor} w_j T^{-1/2}\textstyle\sum_{i=\lfloor\tau_{2,0}T\rfloor+1}^{j-1}(1-\delta_{2,T})^{j-i-1}\varepsilon_i \\
&= O_p(T^{-1/2})
\end{aligned}$$

uniformly in r and so

$$T^{-1/2}y^*_{\lfloor rT\rfloor}-T^{-1/2}\sum_{j=2}^{\lfloor rT\rfloor}w_j\varepsilon_j\xrightarrow{p}0$$

uniformly in r.

Since w_j is independent $N(0,1)$, we have that, conditional on the original sample,

$$T^{-1/2}\sum_{j=2}^{\lfloor rT\rfloor}w_j\varepsilon_j\sim N\left(0,T^{-1}\sum_{j=2}^{\lfloor rT\rfloor}\varepsilon_j^2\right).$$

Then, because

$$T^{-1}\sum_{j=2}^{\lfloor rT\rfloor}\varepsilon_j^2\xrightarrow{p}\int_0^r\omega(h)^2\,dh \tag{A.2}$$

it holds that

$$\begin{aligned}
&T^{-1/2}\sum_{j=2}^{\lfloor rT\rfloor}w_j\varepsilon_j\xrightarrow{w}_p\overline{\omega}W^\eta(r) \\
=\;&\overline{\omega}K^\eta_{0,0}(r).
\end{aligned}$$

Thus we find that

$$T^{-1/2}y^*_{\lfloor rT\rfloor}\xrightarrow{w}_p\overline{\omega}K^\eta_{0,0}(r).$$

Also, since it is easily shown that $\hat{\sigma}^{*2}_\tau=\lfloor\tau T\rfloor^{-1}\sum_{t=1}^{\lfloor\tau T\rfloor}\Delta y^{*2}_t+o_p(1)$ and $\Delta y^*_t=w_t\varepsilon_t+o_p(1)$,

$$\begin{aligned}
\hat{\sigma}^{*2}_\tau &= \lfloor\tau T\rfloor^{-1}\sum_{t=1}^{\lfloor\tau T\rfloor}(w_t\varepsilon_t)^2+o_p(1) \\
&\xrightarrow{p}\tau^{-1}\int_0^\tau\omega(h)^2\,dh.
\end{aligned}$$

It then follows by the CMT that $DF_\tau^* \xrightarrow{w}_p \varpi L_{0,0}^\eta(\tau)$ and that $PWY^* \xrightarrow{w}_p S_{0,0}^\eta$.

A.3 Proof of Theorem 3

As the result simply relies on establishing stochastic orders of magnitude, we give the proof without sub-sample demeaning. The pattern of heteroskedasticity present in ε_t has no effect on these orders.

We first consider the behaviour of DF_τ evaluated at $\tau_{2,0}$. Here

$$DF_{\tau_{2,0}} = \frac{\hat{\phi}_{\tau_{2,0}}}{\sqrt{\hat{\sigma}^2_{\tau_{2,0}} / \sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2}}$$

with

$$\hat{\phi}_{\tau_{2,0}} = \frac{\sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} \Delta y_t y_{t-1}}{\sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2}.$$

Now,

$$\Delta y_t = \begin{cases} \varepsilon_t, & t = 2, ..., \lfloor \tau_{1,0}T \rfloor, \\ \delta_1 y_{t-1} + \varepsilon_t, & t = \lfloor \tau_{1,0}T \rfloor + 1, ..., \lfloor \tau_{2,0}T \rfloor. \end{cases}$$

What follows draws heavily on results established in HLS. In particular, we use the results that $y^2_{\lfloor \tau_{2,0}T \rfloor - 1} = O_p(S_T)$, $\sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2 = O_p(S_T)$ and $\sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} \varepsilon_t y_{t-1} = O_p(S_T^{1/2})$ where $S_T = \lfloor \tau_{1,0}T \rfloor (1+\delta_1)^{2(\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor)}$.

We have

$$\begin{aligned} \sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} \Delta y_t y_{t-1} &= \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} \Delta y_t y_{t-1} + \sum_{\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} \Delta y_t y_{t-1} \\ &= \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_t y_{t-1} + \sum_{\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} (\delta_1 y_{t-1} + \varepsilon_t) y_{t-1} \\ &= \delta_1 \sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2 + \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_t y_{t-1} + \sum_{\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} \varepsilon_t y_{t-1} \end{aligned}$$

so that

$$S_T^{-1} \sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} \Delta y_t y_{t-1} = \delta_1 S_T^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2 + O_p(S_T^{-1/2})$$

and hence

$$\begin{aligned} \hat{\phi}_{\tau_{2,0}} &= \frac{\delta_1 S_T^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2 + O_p(S_T^{-1/2})}{S_T^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} y_{t-1}^2 + S_T^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor + 1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2} \\ &= \delta_1 + O_p(S_T^{-1/2}). \end{aligned}$$

Also,

$$\begin{aligned}
\hat{\sigma}^2_{\tau_{2,0}} &= (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{2,0}T \rfloor} (\Delta y_t - \hat{\phi}_{\tau_{2,0}} y_{t-1})^2 \\
&= (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} (\varepsilon_t - \hat{\phi}_{\tau_{2,0}} y_{t-1})^2 + (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0} \rfloor+1}^{\lfloor \tau_{2,0} \rfloor} \{(\delta_1 - \hat{\phi}_{\tau_{2,0}}) y_{t-1} + \varepsilon_t\}^2.
\end{aligned}$$

Now,

$$\begin{aligned}
(\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} \{(\delta_1 - \hat{\phi}_{\tau_{2,0}}) y_{t-1} + \varepsilon_t\}^2 &= (\delta_1 - \hat{\phi}_{\tau_{2,0}})^2 (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}^2 \\
&\quad + (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} \varepsilon_t^2 \\
&\quad + 2(\delta_1 - \hat{\phi}_{\tau_{2,0}}) (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0}T \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} y_{t-1}\varepsilon_t \\
&= (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=\lfloor \tau_{1,0} \rfloor+1}^{\lfloor \tau_{2,0}T \rfloor} \varepsilon_t^2 + O_p(T^{-1}) \\
&= O_p(1)
\end{aligned}$$

while

$$\begin{aligned}
(\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} (\varepsilon_t - \hat{\phi}_{\tau_{2,0}} y_{t-1})^2 &= (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} \varepsilon_t^2 \\
&\quad + \hat{\phi}^2_{\tau_{2,0}} (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} y_{t-1}^2 \\
&\quad + \hat{\phi}_{\tau_{2,0}} (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} y_{t-1}\varepsilon_t \\
&= \hat{\phi}^2_{\tau_{2,0}} (\lfloor \tau_{2,0}T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau_{1,0}T \rfloor} y_{t-1}^2 + O_p(1) \\
&= O_p(T)
\end{aligned}$$

so $\hat{\sigma}^2_{\tau_{2,0}} = O_p(T)$. This gives

$$DF_{\tau_{2,0}} = \frac{\delta_1 + O_p(S_T^{-1/2})}{\sqrt{O_p(T).O_p(S_T^{-1})}}$$

such that $DF_{\tau_{2,0}}$ diverges to $+\infty$ at a rate $T^{-1/2}S_T^{1/2}$. As a consequence, PWY diverges to $+\infty$ at a rate as least as fast as $T^{-1/2}S_T^{1/2} = \lfloor \tau_{1,0}T^{1/2} \rfloor (1+\delta_1)^{(\lfloor \tau_{2,0}T \rfloor - \lfloor \tau_{1,0}T \rfloor)}$.

Next, it is sufficient to consider the behaviour of DF^*_τ for $\tau \in (\tau_{1,0}, \tau_{2,0}]$. Here,

$$
\begin{aligned}
\sum_{t=2}^{\lfloor \tau T \rfloor} \Delta y_t^* y_{t-1}^* &= \sum_{t=2}^{\lfloor \tau T \rfloor} w_t \Delta y_t \sum_{j=1}^{t-1} w_j \Delta y_j \\
&= \sum_{t=2}^{\lfloor \tau_{1,0} T \rfloor} w_t \Delta y_t \sum_{j=1}^{t-1} w_j \Delta y_j + \sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} w_t \Delta y_t \left(\sum_{j=1}^{\lfloor \tau_{1,0} T \rfloor} w_j \Delta y_j + \sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j \Delta y_j \right) \\
&= \sum_{t=2}^{\lfloor \tau_{1,0} T \rfloor} w_t \varepsilon_t \sum_{j=1}^{t-1} w_j \varepsilon_j + \sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} w_t \Delta y_t \left(\sum_{j=1}^{\lfloor \tau_{1,0} T \rfloor} w_j \varepsilon_j + \sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j \Delta y_j \right) \\
&= \sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} w_t \Delta y_t \sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j \Delta y_j + o_p(.) \\
&= \delta_1^2 \sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} w_t y_{t-1} \sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j y_{j-1} + o_p(.)
\end{aligned}
$$

where $o_p(.)$ notation refers a term which has a smaller order in probablility than the leading term. In a similar fashion we find

$$
\sum_{t=2}^{\lfloor \tau T \rfloor} y_{t-1}^{*2} = \delta_1^2 \sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} \left(\sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j y_{j-1} \right)^2 + o_p(.).
$$

Using the fact that $w_{\lfloor \tau T \rfloor-1}^2 y_{\lfloor \tau T \rfloor-2}^2 = O_p(S_T')$ where $S_T' = \lfloor \tau_{1,0} T \rfloor (1+\delta_1)^{2(\lfloor \tau T \rfloor - \lfloor \tau_{1,0} T \rfloor)}$, we may show that $\sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} (\sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j y_{j-1})^2 = O_p(S_T')$. Likewise, since $w_{\lfloor \tau T \rfloor} w_{\lfloor \tau T \rfloor-1} y_{\lfloor \tau T \rfloor-1} y_{\lfloor \tau T \rfloor-2} = O_p(S_T')$, it can be shown that $\sum_{t=\lfloor \tau_{1,0} T \rfloor+1}^{\lfloor \tau T \rfloor} w_t y_{t-1} \sum_{j=\lfloor \tau_{1,0} T \rfloor+1}^{t-1} w_j y_{j-1} = O_p(S_T')$. Hence,

$$
\begin{aligned}
\hat{\phi}_\tau^* &= \frac{\sum_{t=2}^{\lfloor \tau T \rfloor} \Delta y_t^* y_{t-1}^*}{\sum_{t=2}^{\lfloor \tau T \rfloor} y_{t-1}^{*2}} \\
&= \frac{O_p(S_T')}{O_p(S_T')} \\
&= O_p(1).
\end{aligned}
$$

Also,

$$
\begin{aligned}
\hat{\sigma}_\tau^{*2} &= (\lfloor \tau T \rfloor - 2)^{-1} \sum_{t=2}^{\lfloor \tau T \rfloor} (\Delta y_t - \hat{\phi}_\tau^* y_{t-1})^2 \\
&= (\lfloor \tau T \rfloor - 2)^{-1} (\delta_1 - \hat{\phi}_\tau^*)^2 \sum_{t=2}^{\lfloor \tau T \rfloor} y_{t-1}^{*2} + o_p(.) \\
&= (\lfloor \tau T \rfloor - 2)^{-1} O_p(1) O_p(S_T') + o_p(.) \\
&= O_p(T^{-1} S_T').
\end{aligned}
$$

So,

$$\begin{aligned} DF_\tau^* &= \frac{\hat{\phi}_\tau^*}{\sqrt{\hat{\sigma}_\tau^{*2} / \sum_{t=2}^{\lfloor \tau T \rfloor} y_{t-1}^{*2}}} \\ &= \frac{O_p(1)}{\sqrt{O_p(T^{-1} S_T') O_p(S_T'^{-1})}} \\ &= O_p(T^{1/2}) \end{aligned}$$

and, consequently, $PWY^* = O_p(T^{1/2})$.

References

Balke, N.S. and Wohar, M.E. (2001). Explaining stock price movements: is there a case for fundamentals? Federal Reserve Bank of Dallas Economic and Financial Review, 3rd Quarter, 22–34.

Bettendorf. T. and Chen, W. (2013). Are there bubbles in the sterling-dollar exchange rate? New evidence from sequential ADF tests. Economics Letters 120, 350–353.

Calvo-Gonzalez, O., Shankar, R. and Trezzi, R. (2010). Are commodity prices more volatile now? World Bank Policy Research Working Paper 5460.

Cavaliere, G. and Taylor, A.M.R. (2007). Testing for unit roots in time series models with non-stationary volatility. Journal of Econometrics 140, 919–947.

Cavaliere, G. and Taylor, A.M.R (2008a). Bootstrap unit root tests for time series with non-stationary volatility. Econometric Theory 24, 43–71.

Cavaliere, G. and Taylor, A.M.R (2008b). Time-transformed unit root tests for models with non-stationary volatility. Journal of Time Series Analysis 29, 300–330.

Cavaliere, G. and Taylor, A.M.R. (2009a). Bootstrap M unit root tests. Econometric Reviews 28, 393–421.

Cavaliere, G. and Taylor, A.M.R. (2009b). Heteroskedastic time series with a unit root. Econometric Theory 25, 1228–1270.

Davidson J. (1994). Stochastic Limit Theory. Oxford: Oxford University Press.

Davidson, R. and MacKinnon, J. (2000). Bootstrap tests: how many bootstraps? Econometric Reviews 19, 55–68.

Diba, B.T. and Grossman, H.I. (1988). Explosive rational bubbles in stock prices? American Economic Review 78, 520–530.

Etienne, X. L., Iriwn, S.H. and Garcia, P. (2014). Bubbles in food commodity markets: Four decades of evidence. Journal of International Money and Finance 42, 129–155.

Etienne, X. L., Iriwn, S.H. and Garcia, P. (2015). Price explosiveness, speculation, and grain futures prices. American Journal of Agricultural Economics 97, 65–87.

Ewing, B.T. and Malik, F. (2010). Estimating volatility persistence in oil prices under structural breaks. The Financial Review 45, 1011–1023.

Gilbert, C.L. (2010). Speculative influences on commodity futures prices 2006-2008. Discussion Paper 197, United Nations Conference on Trade and Development (UNCTAD), Geneva.

Gutierrez, L. (2013). Speculative bubbles in agricultural commodity markets. European Review of Agricultural Economics 40, 217–238.

Hansen, B.E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.

Hansen, B.E. (2000). Sample splitting and threshold estimation. Econometrica 68, 575–603.

Harvey, D.I., Leybourne, S.J. and Sollis, R. (2015). Improving the accuracy of asset price bubble start and end date estimators. Discussion Paper, School of Economics, University of Nottingham.

Homm, U. and Breitung, J. (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. Journal of Financial Econometrics 10, 198–231.

Inclan, C. and Tiao, G.C. (1994). Use of cumulative sums of squares for retrospective detection of changes in variance. Journal of the American Statistic Association 89, 913–923.

Kokoszka, P. and Leipus, R. (1999). Testing for parameter changes in ARCH models. Lithuanian Mathematical Journal 39, 182–195.

Liu, R.Y. (1988). Bootstrap procedures under some non i.i.d. models. Annals of Statistics 16, 1696–1708.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. Annals of Statistics 21, 255–285.

McMillan, D.G. and Wohar, M.E. (2011). Structural breaks in volatility: the case of UK sector returns. Applied Financial Economics 21, 578–595.

Phillips, P.C.B. (1987). Towards a unified asymptotic theory for autoregression. Biometrika 74, 535–547.

Phillips, P.C.B. and Shi, S.-P. (2014). Financial bubble implosion. Cowles Foundation Discussion Paper 1967, Yale University.

Phillips, P.C.B., Shi, S.-P. and Yu, J. (2014). Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. International Economic Review, forthcoming.

Phillips, P.C.B., Wu, Y. and Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: when did exuberance escalate stock values? International Economic Review 52, 201–226.

Phillips, P.C.B. and Yu, J. (2010). Dating the timeline of financial bubbles during the subprime crisis. Cowles Foundation Discussion Paper 1770, Yale University.

Phillips, P.C.B. and Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. Quantitative Economics 2, 455–491.

Rapach, D.E., Strauss, J.K. and Wohar, M.E. (2008). Forecasting stock return volatility in the presence of structural breaks. In: Rapach, D.E., Wohar, M.E. (Eds.), Forecasting in the Presence of Structural Breaks and Model Uncertainty. Elsevier\Emerald Series Frontiers of Economics and Globalization, Bingley, UK, pp. 381–416.

Vivian, A. and Wohar, M.E. (2012). Commodity volatility breaks. Journal of International Financial Markets, Institutions and Money 22, 395–422.

Wu, C.F.J. (1986). Jackknife, bootstrap, and other resampling methods. Annals of Statistics 14, 1261–1295.

Table 1. Finite sample powers of nominal 0.05-level tests: $T = 200$, $\tau_{1,0} = 0.4$, $\tau_{2,0} = 0.6$, $\delta_2 = 0$, single volatility shift.

		$\tau_\sigma = 0.3$					$\tau_\sigma = 0.5$					$\tau_\sigma = 0.7$				
σ_1/σ_0	δ_1	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*
1/6	0.00	0.022	0.035	0.053	0.035	0.053	0.034	0.034	0.045	0.034	0.045	0.040	0.033	0.040	0.033	0.040
	0.02	0.047	0.071	0.102	0.080	0.117	0.130	0.130	0.151	0.139	0.157	0.209	0.193	0.210	0.200	0.217
	0.04	0.613	0.630	0.655	0.640	0.658	0.623	0.623	0.638	0.627	0.637	0.632	0.624	0.633	0.626	0.636
	0.06	0.849	0.856	0.864	0.859	0.867	0.852	0.852	0.857	0.856	0.860	0.856	0.853	0.856	0.853	0.858
	0.08	0.936	0.939	0.943	0.941	0.944	0.932	0.932	0.935	0.933	0.935	0.937	0.936	0.937	0.936	0.939
1/3	0.00	0.022	0.035	0.051	0.035	0.051	0.034	0.034	0.045	0.034	0.045	0.040	0.033	0.040	0.033	0.040
	0.02	0.069	0.094	0.126	0.109	0.142	0.136	0.136	0.157	0.146	0.165	0.209	0.193	0.210	0.200	0.218
	0.04	0.618	0.636	0.653	0.643	0.659	0.626	0.626	0.641	0.630	0.638	0.632	0.624	0.633	0.626	0.636
	0.06	0.844	0.853	0.860	0.855	0.862	0.853	0.853	0.857	0.856	0.860	0.856	0.853	0.856	0.853	0.858
	0.08	0.930	0.933	0.938	0.936	0.937	0.933	0.933	0.936	0.933	0.934	0.937	0.936	0.937	0.936	0.938
1	0.00	0.050	0.038	0.042	0.038	0.042	0.050	0.038	0.042	0.038	0.042	0.050	0.038	0.042	0.038	0.042
	0.02	0.219	0.193	0.202	0.194	0.206	0.219	0.193	0.202	0.194	0.206	0.219	0.193	0.202	0.194	0.206
	0.04	0.637	0.622	0.628	0.621	0.628	0.637	0.622	0.628	0.621	0.628	0.637	0.622	0.628	0.621	0.628
	0.06	0.857	0.851	0.854	0.852	0.856	0.857	0.851	0.854	0.852	0.856	0.857	0.851	0.854	0.852	0.856
	0.08	0.937	0.936	0.936	0.936	0.937	0.937	0.936	0.936	0.936	0.937	0.937	0.936	0.936	0.936	0.937
3	0.00	0.387	0.064	0.073	0.064	0.073	0.372	0.065	0.080	0.065	0.080	0.299	0.063	0.087	0.063	0.087
	0.02	0.538	0.162	0.176	0.156	0.177	0.530	0.165	0.187	0.165	0.195	0.435	0.085	0.117	0.093	0.128
	0.04	0.726	0.470	0.483	0.461	0.479	0.746	0.456	0.476	0.456	0.482	0.745	0.471	0.519	0.495	0.538
	0.06	0.865	0.745	0.751	0.741	0.748	0.876	0.733	0.742	0.729	0.746	0.903	0.785	0.804	0.796	0.813
	0.08	0.937	0.888	0.891	0.886	0.889	0.944	0.881	0.886	0.880	0.888	0.958	0.912	0.917	0.914	0.921
6	0.00	0.616	0.078	0.087	0.078	0.087	0.615	0.074	0.083	0.074	0.083	0.571	0.079	0.100	0.079	0.100
	0.02	0.708	0.108	0.122	0.111	0.124	0.712	0.141	0.155	0.144	0.168	0.648	0.068	0.088	0.069	0.093
	0.04	0.818	0.354	0.373	0.346	0.359	0.822	0.306	0.330	0.313	0.344	0.847	0.314	0.352	0.344	0.382
	0.06	0.904	0.678	0.687	0.670	0.674	0.906	0.568	0.585	0.569	0.598	0.945	0.698	0.721	0.715	0.735
	0.08	0.953	0.852	0.856	0.851	0.853	0.950	0.773	0.783	0.770	0.788	0.975	0.879	0.889	0.886	0.893

Table 2. Finite sample powers of nominal 0.05-level tests: $T = 200$, $\tau_{1,0} = 0.4$, $\tau_{2,0} = 0.6$, $\delta_2 = 0$.

		Double volatility shift					Logistic smooth transition in volatility					Trending volatility				
σ_1/σ_0	δ_1	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*	PWY	PWY_1^{adj}	PWY_2^{adj}	PWY^*	PWY_B^*
1/6	0.00	0.047	0.038	0.044	0.038	0.044	0.032	0.034	0.046	0.034	0.046	0.015	0.030	0.041	0.030	0.041
	0.02	0.152	0.131	0.146	0.140	0.153	0.124	0.129	0.149	0.134	0.154	0.116	0.161	0.190	0.173	0.194
	0.04	0.676	0.663	0.673	0.668	0.677	0.626	0.629	0.643	0.633	0.641	0.615	0.648	0.667	0.651	0.660
	0.06	0.867	0.862	0.866	0.866	0.869	0.850	0.851	0.857	0.854	0.860	0.846	0.858	0.866	0.863	0.866
	0.08	0.938	0.936	0.938	0.936	0.937	0.934	0.935	0.936	0.936	0.936	0.932	0.935	0.937	0.935	0.937
1/3	0.00	0.046	0.038	0.044	0.038	0.044	0.032	0.034	0.046	0.034	0.046	0.018	0.031	0.038	0.031	0.038
	0.02	0.158	0.137	0.153	0.149	0.160	0.130	0.133	0.156	0.143	0.162	0.136	0.171	0.188	0.179	0.201
	0.04	0.673	0.659	0.670	0.665	0.673	0.624	0.627	0.642	0.633	0.641	0.619	0.646	0.655	0.648	0.657
	0.06	0.865	0.862	0.865	0.862	0.866	0.852	0.853	0.859	0.854	0.859	0.847	0.857	0.860	0.860	0.863
	0.08	0.939	0.938	0.939	0.938	0.940	0.935	0.935	0.937	0.935	0.936	0.931	0.935	0.936	0.935	0.936
1	0.00	0.050	0.038	0.042	0.038	0.042	0.050	0.038	0.042	0.038	0.042	0.050	0.038	0.042	0.038	0.042
	0.02	0.219	0.193	0.202	0.194	0.206	0.219	0.193	0.202	0.194	0.206	0.219	0.193	0.202	0.194	0.206
	0.04	0.637	0.622	0.628	0.621	0.628	0.637	0.622	0.628	0.621	0.628	0.637	0.622	0.628	0.621	0.628
	0.06	0.857	0.851	0.854	0.852	0.856	0.857	0.851	0.854	0.852	0.856	0.857	0.851	0.854	0.852	0.856
	0.08	0.937	0.936	0.936	0.936	0.937	0.937	0.936	0.936	0.936	0.937	0.937	0.936	0.936	0.936	0.937
3	0.00	0.291	0.068	0.093	0.068	0.093	0.367	0.063	0.075	0.063	0.075	0.224	0.048	0.056	0.048	0.056
	0.02	0.452	0.186	0.230	0.182	0.230	0.530	0.167	0.188	0.170	0.197	0.414	0.179	0.191	0.179	0.193
	0.04	0.655	0.439	0.481	0.439	0.476	0.742	0.465	0.483	0.468	0.500	0.696	0.551	0.562	0.552	0.564
	0.06	0.811	0.696	0.718	0.696	0.717	0.874	0.740	0.750	0.739	0.753	0.865	0.803	0.810	0.800	0.808
	0.08	0.903	0.857	0.867	0.858	0.865	0.941	0.887	0.889	0.886	0.893	0.939	0.911	0.914	0.913	0.916
6	0.00	0.530	0.077	0.101	0.077	0.101	0.620	0.075	0.088	0.075	0.088	0.389	0.052	0.060	0.052	0.060
	0.02	0.640	0.145	0.184	0.150	0.190	0.717	0.152	0.169	0.152	0.171	0.538	0.171	0.182	0.168	0.181
	0.04	0.754	0.321	0.363	0.319	0.358	0.828	0.340	0.363	0.338	0.367	0.742	0.512	0.520	0.504	0.521
	0.06	0.846	0.572	0.604	0.569	0.593	0.904	0.587	0.605	0.588	0.614	0.882	0.773	0.778	0.771	0.779
	0.08	0.912	0.777	0.794	0.777	0.789	0.951	0.792	0.802	0.791	0.804	0.947	0.901	0.903	0.899	0.904

Table 3. Tests for stationary volatility in commodity prices, 2000-2013

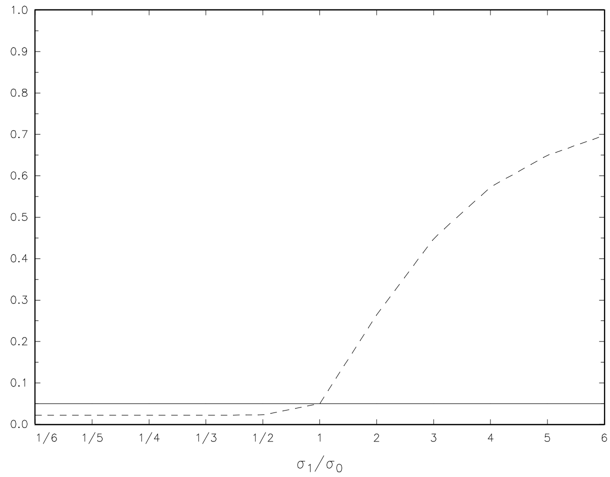
	$\mathcal{H}_{KS}$	$\mathcal{H}_{R}$	$\mathcal{H}_{CVM}$	$\mathcal{H}_{AD}$
		Monthly real series		
Brent oil	0.978	1.090	0.339	2.277*
WTI oil	1.209	1.309	0.496**	3.276**
Gold	1.751***	1.751**	1.037***	6.852***
Silver	1.578**	1.589	1.088***	6.946***
Platinum	1.341*	1.341	0.585**	3.869***
Aluminium	1.025	1.215	0.325	2.170*
Copper	1.368**	1.504	0.641**	4.206***
		Weekly nominal series		
Brent oil	2.228***	2.405***	1.835***	12.389***
WTI oil	1.892***	2.196***	1.057***	7.264***
Gold	2.536***	2.536***	2.647***	17.240***
Silver	2.029***	2.092***	1.795***	11.653***
Platinum	2.219***	2.333***	1.334***	9.152***
Aluminium	2.167***	2.536***	1.290***	8.476***
Copper	2.842***	3.210***	2.183***	14.533***

Notes: *, ** and *** denote rejection at the 0.10-, 0.05- and 0.01-level, respectively. Critical values for the $\mathcal{H}_{KS}$, $\mathcal{H}_{R}$, $\mathcal{H}_{CVM}$ and $\mathcal{H}_{AD}$ tests are given in Shorack and Wellner (1987): Table 1, p. 413; Table 2, p. 144; Table 4, p. 147 and Table 5, p. 148; respectively.

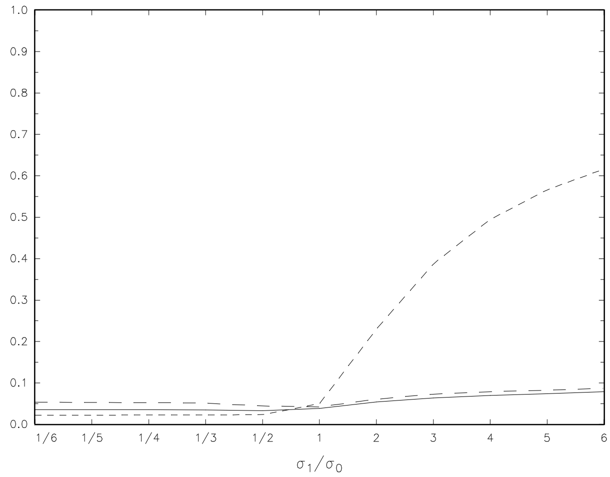
Table 4. Tests for a bubble in commodity prices, 2000-2013

		p-values		
	PWY statistic	PWY	PWY^*	PWY^*_B
		Monthly real series		
Brent oil	2.073	0.008	0.097	0.034
WTI oil	2.230	0.006	0.105	0.043
Gold	3.306	0.000	0.095	0.074
Silver	4.809	0.000	0.067	0.026
Platinum	2.547	0.001	0.167	0.092
Aluminium	2.602	0.001	0.055	0.020
Copper	5.901	0.000	0.006	0.001
		Weekly nominal series		
Brent oil	3.112	0.000	0.043	0.024
WTI oil	2.986	0.000	0.093	0.059
Gold	4.223	0.000	0.056	0.052
Silver	5.899	0.000	0.024	0.019
Platinum	3.887	0.000	0.071	0.053
Aluminium	2.659	0.000	0.082	0.069
Copper	7.748	0.000	0.016	0.009

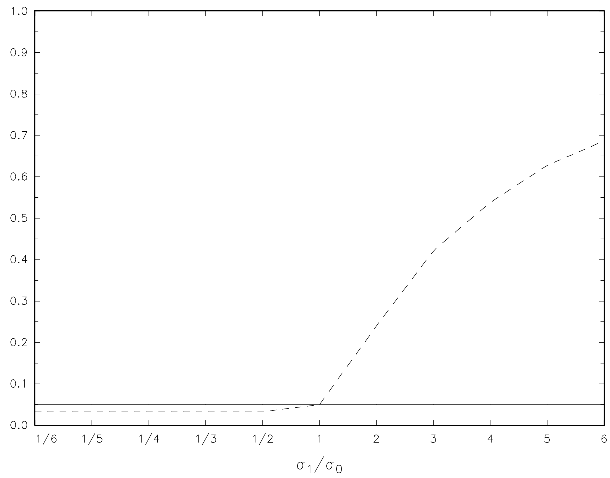
Note: p-values for PWY were obtained by simulating the finite sample distribution of PWY for $T = 168$ and $T = 731$ for monthly and weekly data respectively, using 5,000 replications of a random walk with homoskedastic $IIDN(0,1)$ innovations.



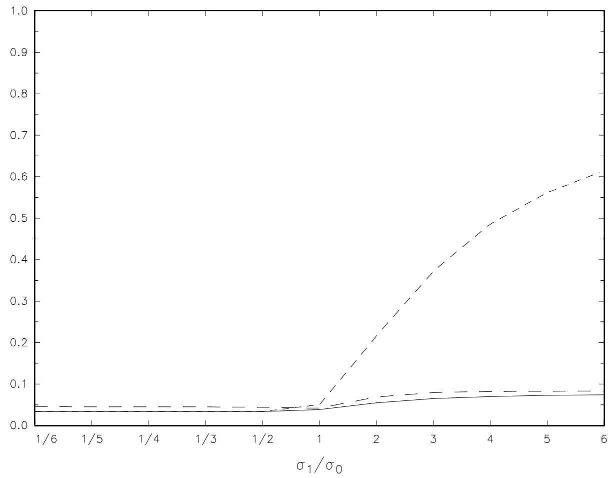
(a) $\tau_\sigma = 0.3$, $T = \infty$



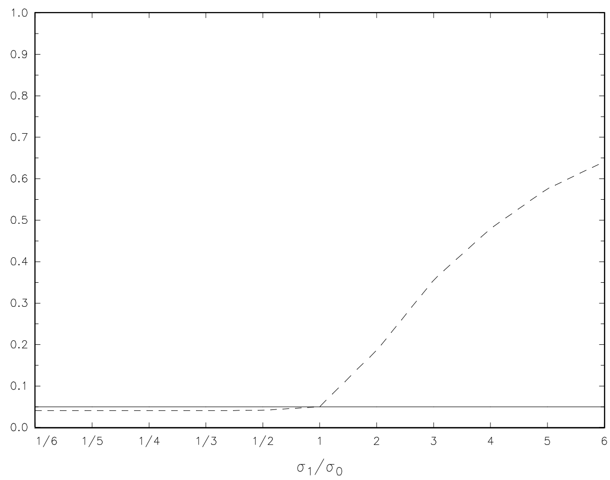
(b) $\tau_\sigma = 0.3$, $T = 200$



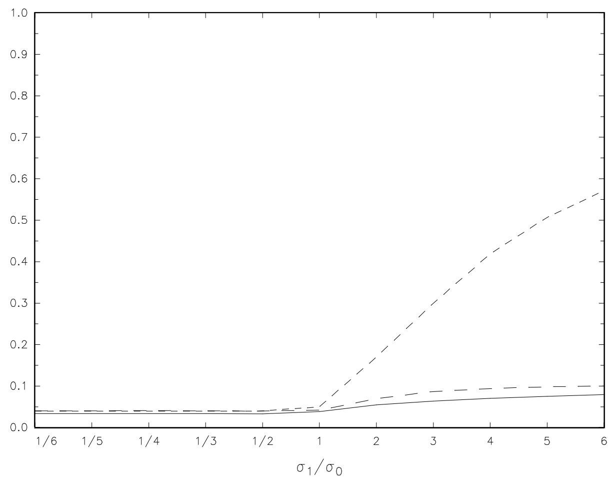
(c) $\tau_\sigma = 0.5$, $T = \infty$



(d) $\tau_\sigma = 0.5$, $T = 200$

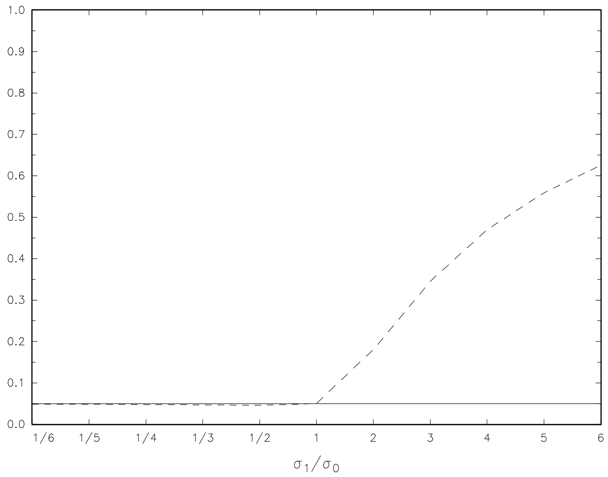


(e) $\tau_\sigma = 0.7$, $T = \infty$

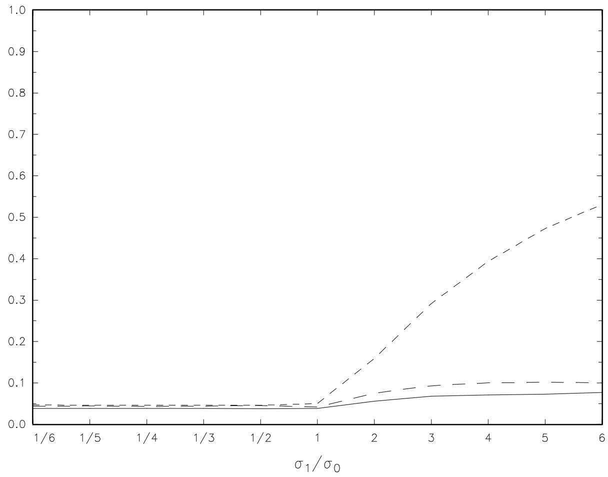


(f) $\tau_\sigma = 0.7$, $T = 200$

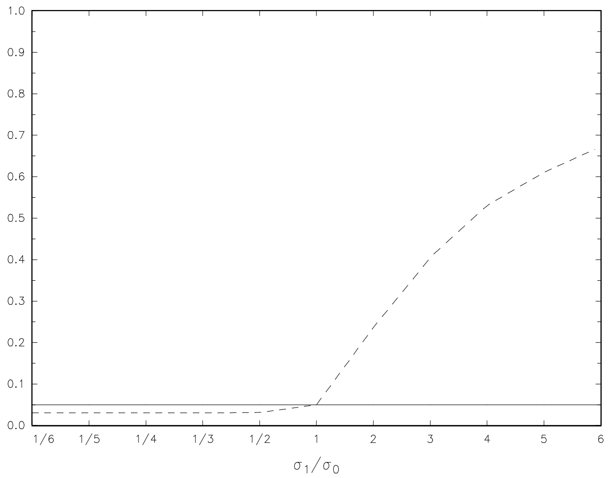
Figure 1. Asymptotic and finite sample size of nominal 0.05-level tests: single volatility shift; PWY: - - - , PWY^*: —— , PWY_B^*: – –



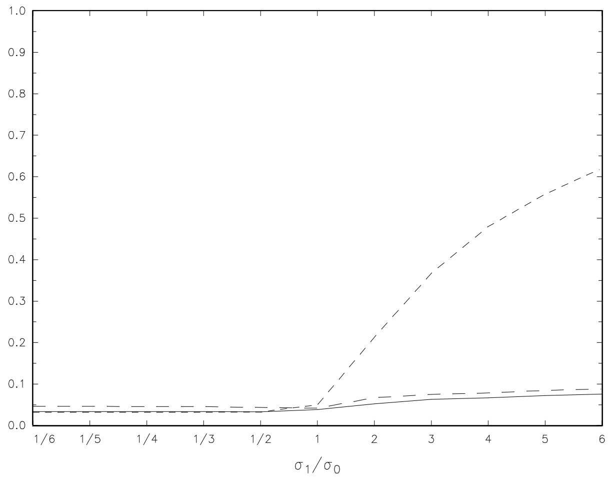
(a) Double volatility shift, $T = \infty$



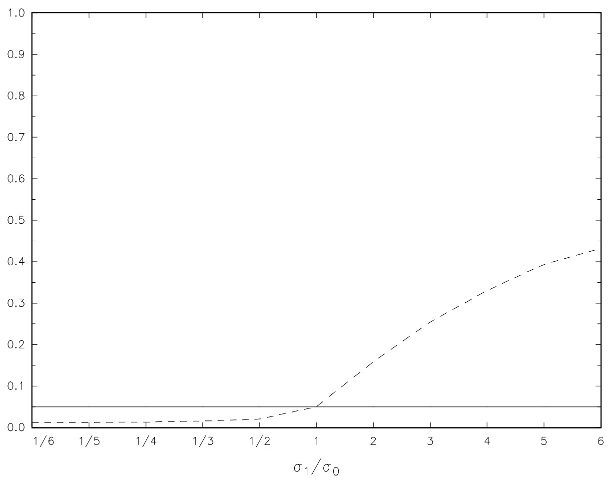
(b) Double volatility shift, $T = 200$



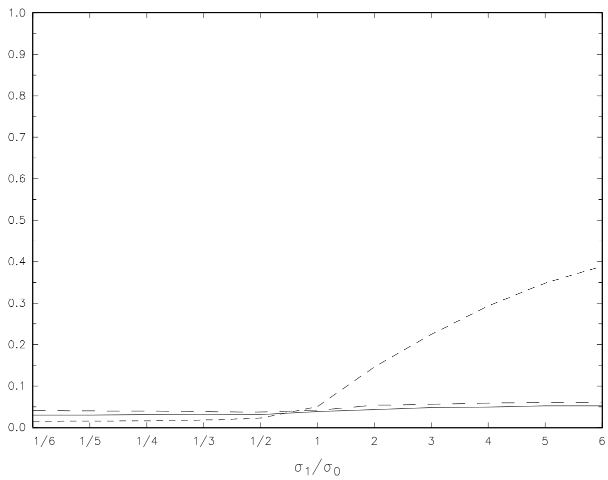
(c) Logistic smooth transition in volatility, $T = \infty$



(d) Logistic smooth transition in volatility, $T = 200$

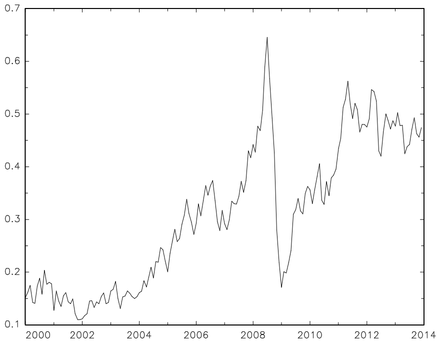


(e) Trending volatility, $T = \infty$

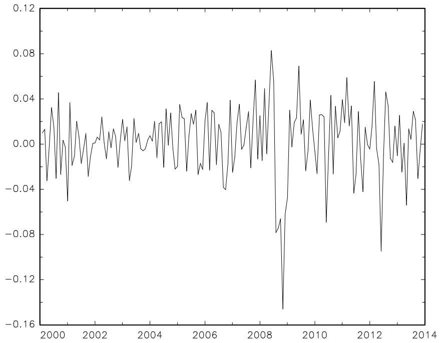


(f) Trending volatility, $T = 200$

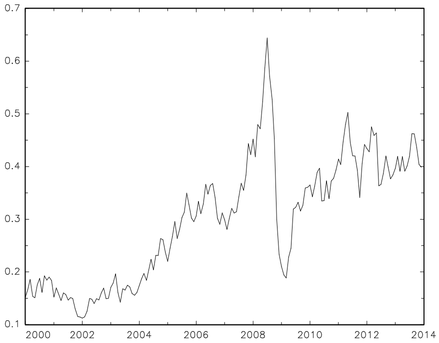
Figure 2. Asymptotic and finite sample size of nominal 0.05-level tests:
PWY: - - - , PWY^*: —— , PWY^*_B: – –



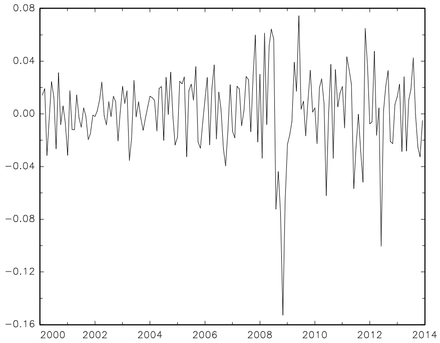
Brent oil: levels



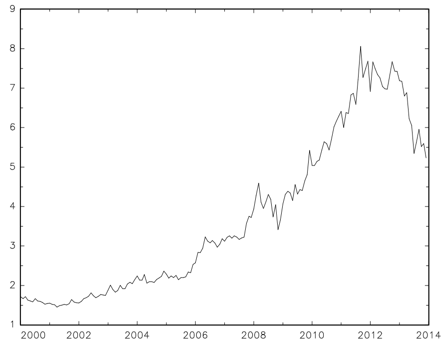
Brent oil: first differences



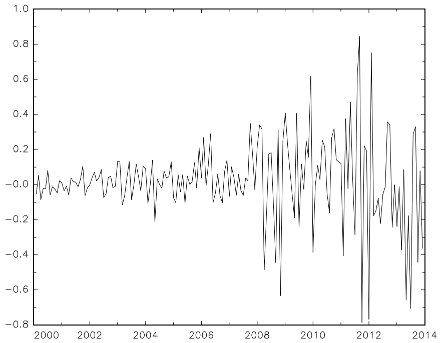
WTI oil: levels



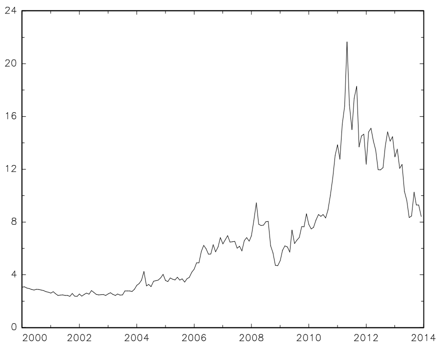
WTI oil: first differences



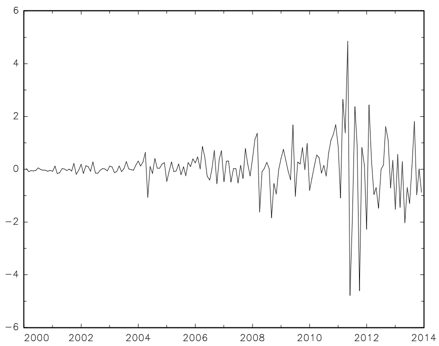
Gold: levels



Gold: first differences

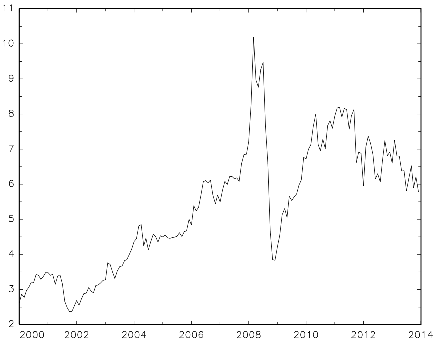


Silver: levels

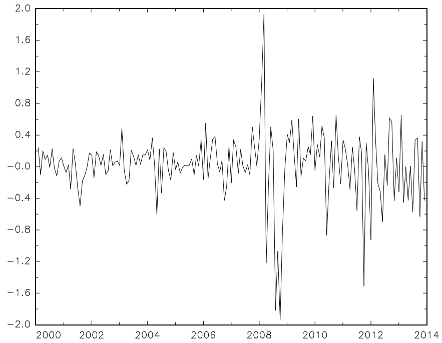


Silver: first differences

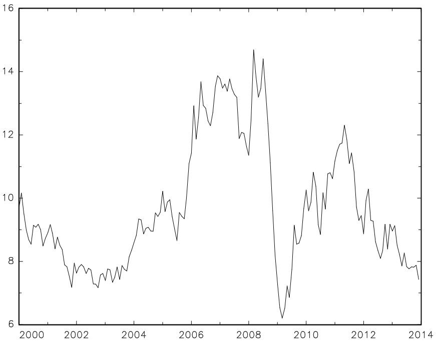
Figure 3(a). Monthly real commodity prices, 2000-2013



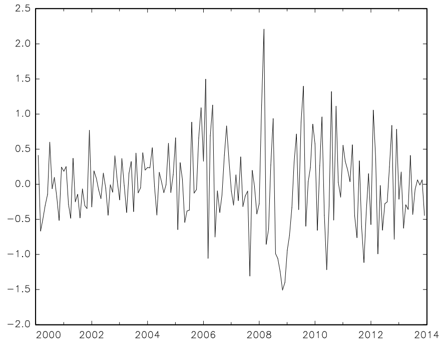
Platinum: levels



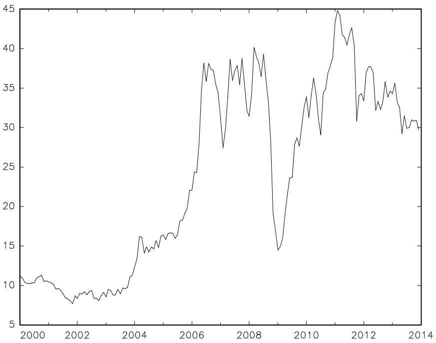
Platinum: first differences



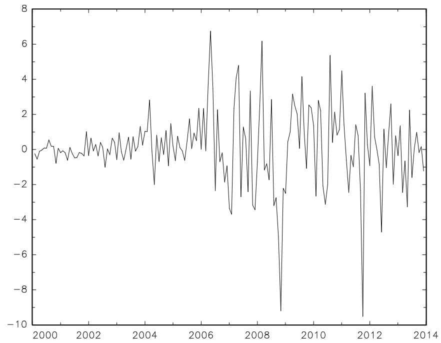
Aluminium: levels



Aluminium: first differences

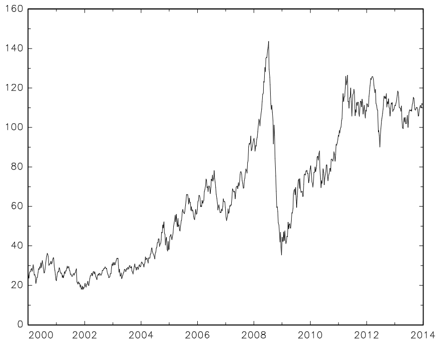


Copper: levels

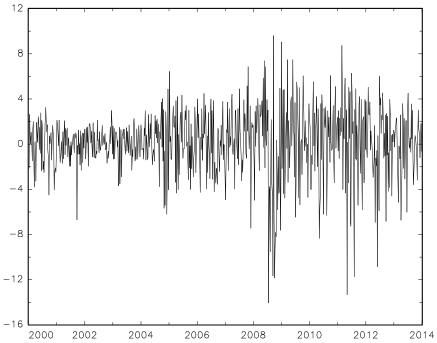


Copper: first differences

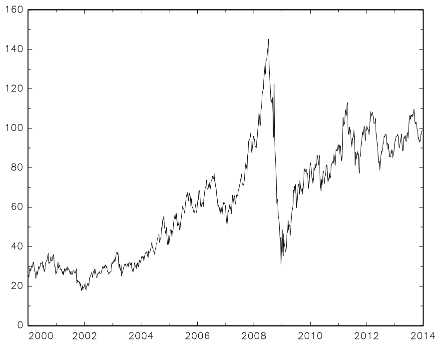
Figure 3(b). Monthly real commodity prices, 2000-2013



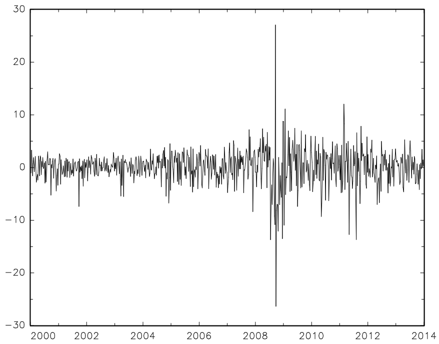
Brent oil: levels



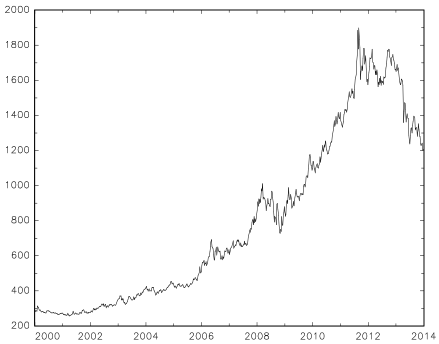
Brent oil: first differences



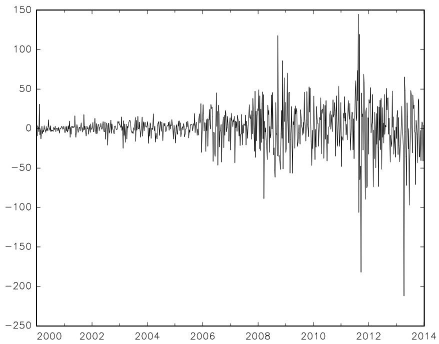
WTI oil: levels



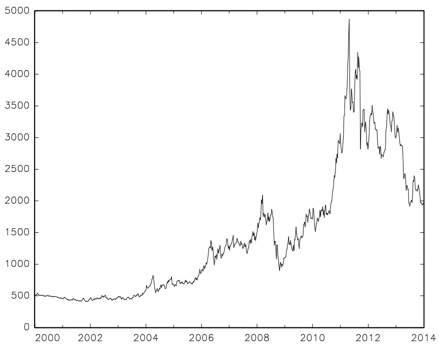
WTI oil: first differences



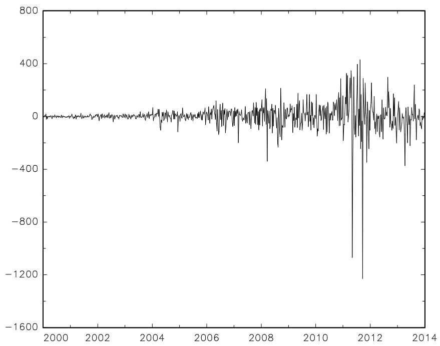
Gold: levels



Gold: first differences

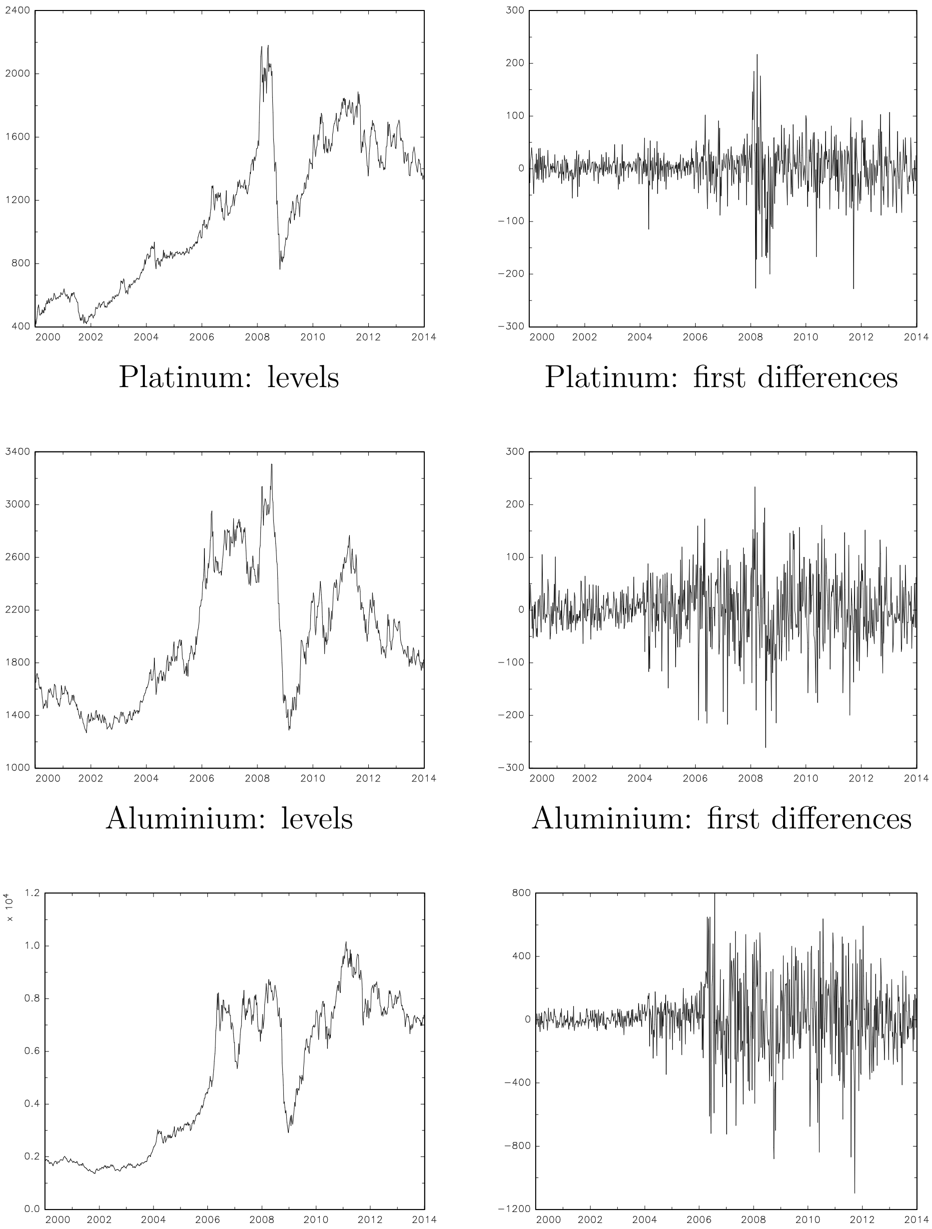


Silver: levels



Silver: first differences

Figure 4(a). Weekly nominal commodity prices, 2000-2013



Platinum: levels

Platinum: first differences

Aluminium: levels

Aluminium: first differences

Copper: levels

Copper: first differences

Figure 4(b). Weekly nominal commodity prices, 2000-2013